VOLUME ONE

ENCYCLOPEDIA OF STRESS

SECOND EDITION

EDITOR-IN-CHIEF
GEORGE FINK

SHOUT
SHARE
IT

DEDICATION

For Ann Elizabeth

EDITOR-IN-CHIEF

George Fink
Professorial Research Fellow (formerly Director)
Mental Health Research Institute of Victoria
Parkville, Melbourne, Victoria
Australia
Formerly Director, MRC Brain Metabolism Unit
Edinburgh, Scotland, UK.

ASSOCIATE EDITORS

Bruce McEwen
The Rockefeller University
Laboratory of Neuroendocrinology
1230 York Avenue, Box 165
New York NY 10021

E. Ronald de Kloet
Universiteit Leiden
Natuurwetenschappen Wiskunde en
LACDR Medical Pharmacology
Einsteinweg 55, kamer HB 902
Leiden 2333 CC
The Netherlands

Robert Rubin
Department of Psychiatry (116A)
VA Greater LA Healthcare System
11301 Wilshire Blvd.
Los Angeles, CA 90073

George Chrousos
First Department of Pediatrics,
Athens University Medical School,
Aghia Sophia Children's Hospital,
115 27 Athens, Greece

Andrew Steptoe
University College London
Department of Epidemiology and Public Health
Psychology
1-19 Torrington Place London WC1E 6BT
UK

Noel Rose
Johns Hopkins University
Center for Autoimmune Disease Research
Molecular Microbiology and Immunology
615 North Wolfe Street, Rm E5014
Baltimore MD 21205

Ian Craig
SGDP Research Centre
Institute of Psychiatry, Molecular Genetics Group
De Crespigny Park, Box P082
Denmark Hill
London SE5 8AF
UK

Giora Feuerstein
Merck Research Laboratories 42-209
Department of Cardiovascular Diseases
770 Sumneytown Pike
West Point PA 19486

CONTENTS

Contents by Subject area *xxiii*
Preface to First Edition *xxxi*
Preface to the Second Edition *xxxiii*
Guide to Encyclopedia *xxxv*
Foreword *xxxvii*

VOLUME 1

A

Acute Stress Disorder and Posttraumatic Stress Disorder **1**
R. Yehuda and C. M. Wong

Acute Stress Response: Experimental **7**
K. Pacak and R. McCarty

Acute Trauma Response **15**
W. C. Chiu, D. E. Carlson and M. P. Lilly

Adenylyl Cyclases and Stress Response **21**
F. A. Antoni

Adjustment Disorders **24**
M. Dascalu and D. Svrakic

Adolescence **28**
G. N. Swanson

Adolescent Suicide **36**
M. Berk, R. Suddath and M. Devich-Navarro

Adrenal Cortex **38**
G. P. Vinson, B. J. Whitehouse and J. P. Hinson

Adrenal Insufficiency **47**
H. S. Willenberg, S. R. Bornstein and G. P. Chrousos

Adrenal Medulla **52**
R. Kvetnansky and R. McCarty

Adrenaline **60**
T. M. Pollard

Adrenocortical Function, Factors Controlling Development Thereof **64**
T. Else and G. D. Hammer

Adrenocorticotropic Hormone (ACTH) **69**
M. E. Rhodes

Aerobic Exercise and Stress Reduction **73**
E. J. C. de Geus and J. H. Stubbe

Affective Disorders **78**
D. F. MacKinnon

Aggression **84**
E. F. Coccaro and E. C. Manning

Aggressive Behavior **89**
J. M. Koolhaas

Aging and Adrenocortical Factors **92**
C. Lord and J. C. Pruessner

Aging and Psychological Stress **96**
B. W. J. Penninx

Aging and Stress, Biology of **102**
M. A. Horan, R. N. Barton and G. J. Lithgow

AIDS **108**
M. H. Antoni and D. G. Cruess

Airline Accidents **114**
G. Li

Alarm Phase and General Adaptation Syndrome **119**
R. McCarty and K. Pacak

Alcohol and Stress: Social and Psychological Aspects 123
M. A. Sayette

Alcohol, Alcoholism, and Stress: A Psychobiological Perspective 127
A. N. Taylor, P. Prolo and M. L. Pilati

Aldosterone and Mineralocorticoid Receptors 132
J. W. Funder

Allostasis and Allostatic Load 135
B. S. McEwen and J. C. Wingfield

Alzheimer's Disease 141
A. E. Roth, W. M. Greenberg and N. Pomara

Ambulatory Blood Pressure Monitoring 147
T. G. Pickering

Amenorrhea 149
V. E. Beshay and B. R. Carr

Amnesia 158
J. S. Simons and K. S. Graham

Amygdala 161
M. S. Fanselow and R. Ponnusamy

Anatomy of the HPA Axis 166
A. G. Watts

Androgen Action 171
R. J. Handa and T. R. Pak

Anger 176
R. W. Novaco

Angiotensin 182
O. Baltatu and M. Bader

Angiotensin Receptors 186
T. A. Jenkins and F. A. O. Mendelsohn

Animal Models (Nonprimate) for Human Stress 190
J. E. Ottenweller

Annexin 195
A. Mulla, J. C. Buckingham and R. J. Flower

Antibody Response 199
R. J. Booth

Anti-CRF 206
E. P. Zorrilla, Y. Zhao and G. F. Koob

Antidepressant Actions on Glucocorticoid Receptors 215
J. L. W. Yau and J. R. Seckl

Antimineralocorticoid Challenge 223
M. Kellner and K. Wiedemann

Antipsychotic Drugs and Stress 226
S. Sundram

Antisocial Disorders 231
K. Pajer

Anxiety 236
A. Öhman

Anxiolytics 240
M. Lader

Apoptosis 244
J.-L. Turner and J. A. Cidlowski

Arterial Baroreflex 248
G. Parati, P. Castiglioni, M. Di Rienzo and G. Mancia

Arthritis 258
R. L. Wilder

Arthritis – Psychological 261
J. W. Younger and A. J. Zautra

Asthma 264
A. A. Kaptein

Atherosclerosis 267
M. W. Ketterer

Attention-Deficit/Hyperactivity Disorder, Stress and 271
L. E. Arnold and R. L. Lindsay

Autoimmunity 276
B. S. Rabin

Autonomic Nervous System 282
W. R. Lovallo and J. J. Sollers III

Autotolerance 290
M. Durai and K. D. Moudgil

Avoidance 297
M. S. Oitzl

B

Behavior Therapy 303
P. de Silva

Behavior, Overview 307
R. Dantzer

Benzodiazepines 310
J. Bermak, T. Johnstone and K. Gee

Bereavement 317
P. J. Clayton

Beta-Adrenergic Blockers 323
M. B. Hamner and G. W. Arana

Beta-Endorphin 332
M. Lee and S. L. Wardlaw

Blood Pressure 335
A. Sherwood and R. A. Carels

Blood–Brain Barrier, Stress and 342
S. N. Malaeb and B. S. Stonestreet

Borderline Personality Disorder 348
H. W. Koenigsberg and L. J. Siever

Brain and Brain Regions 351
A. G. Watts

Brain Natriuretic Peptide (BNP) 357
P. Pervanidou and G. P. Chrousos

Brain Trauma 360
B. Pentland

Breast Cancer 364
A. Moyer

Burnout 368
C. Maslach and M. P. Leiter

C

Calbindin 373
F. A. Antoni

Calcium, Role of 374
F. A. Antoni

Calcium-Dependent Neurotoxicity 375
J. S. Kelly

Cancer 378
D. Spiegel

Cancer Treatment 384
F. I. Fawzy, A. L. Canada and N. W. Fawzy

Captivity, Adaptation to 388
R. H. Rahe

Captivity, Recovery from 392
R. H. Rahe

Cardiovascular System and Stress 396
P. Hjemdahl

Cardiovascular Disease, Stress and 410
G. P. Chrousos and G. Kaltsas

Caregivers, Stress and 416
S. H. Zarit, K. Bottigi and J. E. Gaugler

Catecholamines 419
U. Lundberg

Central Stress Neurocircuits 424
S. Kollack-Walker, H. E. W. Day and H. Akil

Cerebral Metabolism, Brain Imaging 432
K. P. Ebmeier, C. L. Donaghey and N. J. Dougall

Chaperone Proteins and Chaperonopathies 438
A. J. L. Macario and E. Conway de Macario

Chaperonopathies 444
A. J. L. Macario and E. Conway de Macario

Chemical Warfare 449
J. Berberich

Chernobyl, Stress Effects of 452
A. Tønnessen and L. Weisæth

Child Abuse 457
C. C. Swenson and L. Saldana

Child Physical Abuse 460
C. C. Swenson

Child Sexual Abuse 463
J. A. Cohen

Childbirth and Stress 467
S. Ayers and E. Ford

Childhood Stress 472
S. Sandberg

Cholesterol and Lipoproteins 478
C. M. Stoney

Chronic Fatigue Syndrome 484
A. J. Cleare and S. Wessely

Chronic Social Stress: GR Sensitivity in Leukocytes 493
A. Weizman and B. Rotberg

Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress 496
M. P. Antoch and R. V. Kondratov

Circadian Rhythm Effects on Cardiovascular and Other Stress-Related Events 500
R. Manfredini, B. Boari, R. Salmi, A. M. Malagoni and F. Manfredini

Circadian Rhythms, Effects of Prenatal Stress in Rodents 505
S. Maccari and O. Van Reeth

Circadian Rhythms, Genetics of 508
F. W. Turek and M. H. Vitaterna

Cognition and Stress 513
M. W. Eysenck

Cognitive-Behavioral Therapy 515
G. A. Fava

Combat Reaction, Chronic 518
R. H. Rahe

Combat Stress Reaction 524
M. Dobson

Combat, Acute Reactions to 529
R. H. Rahe

Common Cold and Stress 533
A. Smith

Community Studies 536
C. J. Holahan, R. H. Moos and L. M. Groesz

Comorbid Disorders and Stress 542
S. Sundram and Avril Pereira

Comparative Anatomy and Physiology 549
A. G. Watts

Concentration Camp Survivors 553
J. D. Kinzie

Congenital Adrenal Hyperplasia (CAH) 556
A. Solomon and P.-M. G. Bouloux

Conservation of Resources Theory 562
S. E. Hobfoll and J. S. Ford

Control and Stress 568
A. Steptoe

Coping and Stress: A Lens and Filter Model 574
R. H. Rahe

Coping Skills 578
A. DeLongis and E. Puterman

Corticosteroid Receptor Genes: Functional Dissection in Mice 584
F. Tronche

Corticosteroid Receptors 594
O. C. Meijer, E. R. de Kloet and B. S. McEwen

Corticosteroid-Binding Globulin (Transcortin) 605
B. E. P. Murphy

Corticosteroids and Stress 613
A. Munck

Corticotropin Releasing Hormone (CRH) 620
A. T. Lim

Corticotropin Releasing Factor Receptor Deficiency in Mice 623
N. J. Justice and K.-F. Lee

Corticotropin Releasing Factor-Binding Protein 627
P. J. Lowry, C. F. Kemp and R. J. Woods

Corticotropin-Releasing Factor Circuitry in the Brain – Relevance for Affective Disorders and Anxiety 630
D. A. Gutman and C. B. Nemeroff

Corticotropin-Releasing Factor (CRF) Family of Neuropeptides – Role in Inflammation 635
A. Gravanis and A. N. Margioris

Corticotropin-Releasing Factor Receptors 641
D. E. Grigoriadis

Cortisol Awakening Response 649
A. Steptoe

C-Reactive Protein 653
W. J. Kop and A. A. Weinstein

Crime Victims 659
I. Robbins

Crisis Intervention 662
D. Hamaoka, D. Benedek, T. Grieger and R. J. Ursano

Critical Thermal Limits 667
J. Roth

Crowding Stress 669
L. Kovács and P. Csermely

Cultural Factors in Stress 672
J. W. Berry and B. Ataca

Cultural Transition 678
M. S. Kopp

Cushing's Syndrome, Medical Aspects 682
S. R. Bornstein, M. Gruber, H. S. Willenberg, C. A. Stratakis and G. P. Chrousos

Cushing's Syndrome, Neuropsychiatric Aspects 688
M. N. Starkman

Cytokines 692
G. D. Marshall

Cytokines, Chronic Stress, and Fatigue 698
S. Jain and P. J. Mills

Cytokines, Stress, and Depression 705
B. E. Leonard and C. Song

Cytotoxic Lymphocytes 711
M. A. Fletcher and N. G. Klimas

D

Death Anxiety 717
R. Kastenbaum

Defensive Behaviors 722
D. C. Blanchard, M. Yang, M. Hebert and R. J. Blanchard

Demand–Control Model 727
T. Theorell

Dental Stress 733
T. K. Fábián, G. Fábián and P. Fejérdy

Depersonalization: Systematic Assessment 736
M. Steinberg

Depression and Coronary Heart Disease 741
F. Lespérance and N. Frasure-Smith

Depression and Manic–Depressive Illness 744
R. T. Rubin and B. J. Carroll

Depression and Stress, Role of n-3 and n-6 Fatty Acids 754
C. Song and B. E. Leonard

Depression Models 760
K. Matthews and C. Stewart

Depression, Immunological Aspects 766
M. R. Irwin

Dermatological Conditions 773
M. A. Gupta

Desensitization 778
F. A. Antoni

Dexamethasone Suppression Test (DST) 780
R. T. Rubin and B. J. Carroll

DEX-CRH Test 784
N. C. Schommer and I. Heuser

DHEA 788
J. Herbert

Diabetes, Type 1 792
A. Riazi and C. Bradley

Diet and Stress, Non-Psychiatric 797
J. Wardle and E. L. Gibson

Diet and Stress, Psychiatric 806
V. March and M. H. Fernstrom

Disaster Syndrome 811
P. Valent

Disasters and Mass Violence, Public, Effects of 814
G. Stevens, B. Raphael and M. Dobson

Disease, Stress Induced 824
H. S. Willenberg, S. R. Bornstein and G. P. Chrousos

Dissociation 828
J. R. Maldonado

Distress 838
G. Matthews

Divorce, Children of 844
K. N. Hipke, S. A. Wolchik and I. N. Sandler

Domestic Violence 848
B. Donohue, H. Hill and T. Maier-Paarlberg

Dopamine, Central 852
G. D. Stanwood

***Drosophila* Genes and Anoxia** 859
G. G. Haddad

***Drosophila* Studies** 864
M. Allikian and J. Tower

Drug Use and Abuse 866
J. R. Mantsch

E

Earthquakes, Stress Effects of 871
M. Livanou and M. Başoğlu

Eating Disorders and Stress 876
D. C. Jimerson

Eclampsia and Pre-Eclampsia 880
A. Makrigiannakis, G. Petsas and G. P. Chrousos

Economic Factors and Stress 884
R. A. Catalano

Education Levels and Stress 888
J. Mirowsky and C. E. Ross

Effort–Reward Imbalance Model 893
J. Siegrist

Elder Abuse 896
C. P. Holstege and H. Holstege

Electrodermal Activity 899
G. Turpin and T. Grandfield

Emergency Personnel, Stress in 903
D. S. Weiss

Emotional Inhibition 908
H. C. Traue and R. M. Deighton

Emotions: Structure and Adaptive Functions 914
R. J. Contrada and H. Leventhal

Employee Assistance and Counseling 922
M. E. Mor Barak and D. J. Travis

Endocrine Systems 926
G. Fink

Endometriosis 927
S. N. Kalantaridou, T. Vrekoussis, A. Makrigiannakis and G. P. Chrousos

Enuresis 931
S. K. Anand and C. D. Berkowitz

Environmental Factors 934
W. R. Avison

Environmental Stress, Effects on Human Performance 940
G. R. J. Hockey

Epilepsy 945
C. J. Schramke and K. M. Kelly

Estrogen 951
S. B. Miller, E. Neumark and A. Sita

Ethanol and Endogenous Opioids 955
D. K. Sarkar

Ethnicity, Mental Health 959
K. Iley and J. Y. Nazroo

Evolutionary Origins and Functions of the Stress Response 964
R. M. Nesse, S. Bhatnagar and E. A. Young

Excitatory Amino Acids 970
J. V. Nadler

Excitotoxins 975
M. P. Mattson

Exercise 982
P. Khatri and J. A. Blumenthal

Expression Profiling of Stress Responsive Gene Patterns 986
N. A. Datson and M. C. Morsink

VOLUME 2

F

Familial Patterns of Stress 1
A. Bifulco

Family Therapy 6
B. Jalali

Fatigue and Stress 11
A. Appels and W. J. Kop

Fear 15
A. Öhman

Fear and the Amygdala 19
R. Norbury and G. M. Goodwin

Febrile Response 24
S. Gatti McArthur and T. Bartfai

Feedback Systems 31
G. Fink

Feeding Circuitry (and Neurochemistry) 43
S. E. La Fleur, J. J. G. Hillebrand and R. A. H. Adan

Fetal Stress 46
M. Eleftheriades, P. Pervanidou and G. P. Chrousos

Fibrinogen and Clotting Factors 51
E. Brunner

Fibromyalgia 56
P. B. Wood

Fight-or-Flight Response 62
R. McCarty

Firefighters, Stress in 64
T. L. Guidotti

Fish, Stress in 68
C. B. Schreck

Floods, Stress Effects of 73
J. O. Brende

Food Intake and Stress, Human 79
K. Smith and G. M. Goodwin

Food Intake and Stress, Non-Human 82
Belinda A. Henry and Iain J. Clarke

Food Shift Effect 88
F. K. Stephan

Freud, Sigmund 90
D. J. Lynn

G

GABA (Gamma Aminobutyric Acid) 97
J. D. C. Lambert

Gastrointestinal Effects 109
R. Murison and A. M. Milde

Gender and Stress 115
R. J. Handa and W. C. J. Chung

Gene–Environment Interactions in Early Development 122
I. S. P. Davis and R. Plomin

Genetic Factors and Stress 128
C. A. Koch and C. A. Stratakis

Genetic Polymorphisms in Stress Response 135
I. W. Craig

Genetic Predispositions to Stressful Conditions 141
D. Blackwood and H. Knight

Genetic Testing and Stress 146
V. Senior and M. Cropley

Genetic Variation of HPA Axis Activity and Function in Farm Animals 150
P. Mormède

Ghrelin and Stress Protection 153
T. Brzozowski, M. Pawlik, D. Drozdowicz, Z. Sliwowski, S. J. Konturek, W. W. Pawlik and P. C. Konturek

Glia or Neuroglia 161
G. W. Bennett and D. E. Ray

Glucocorticoid Effects on Memory: the Positive and the Negative 166
O. T. Wolf

Glucocorticoid Negative Feedback 172
M. F. Dallman

Glucocorticoid Receptor Mutant Mice as Models for Stress-Induced Affective Disorders 176
P. Gass

Glucocorticoid Receptor Mutations and Polymorphisms 183
J. W. Koper

Glucocorticoids – Adverse Effects on the Nervous System 185
R. M. Sapolsky

Glucocorticoids, Effects of Stress on 190
J. C. Buckingham

Glucocorticoids, Overview 198
B. E. Pearson Murphy

Glucocorticoids, Role in Stress 210
J. C. Buckingham

Glucose Transport 217
A. L. McCall

Glycobiology of Stress 222
G. Lauc and M. Flögel

Gonadotropin Secretion, Effects of Stress on 228
M. Ferin

Graves' Disease (Thyrotoxicosis) 234
W. M. Wiersinga

Grieving 238
A. Ray and H. Prigerson

Group Therapy 242
N. Wong

Gulf War Syndrome, Psychological and Chemical Stressors 248
H. Soreq

H

Health and Socioeconomic Status 255
T. Chandola and M. Marmot

Health Behavior and Stress 262
A. Steptoe

Heart Disease/Attack 266
G. J. Baker, S. Suchday and D. S. Krantz

Heart Failure, Stress Effects 272
M. Alevizaki and G. P. Chrousos

Heart Rate 274
J. R. Jennings

Heat Resistance 278
A. J. L. Macario and E. Conway de Macario

Heat Shock Genes, Human 284
A. J. L. Macario and E. Conway de Macario

Heat Shock Proteins: HSP60 Family Genes 288
H. Kubota

Heat Shock Response, Overview 292
A. J. L. Macario and E. Conway de Macario

Hemostasis and Stress 300
Roland von Känel

Herpesviruses 305
D. A. Padgett, M. T. Bailey and J. F. Sheridan

Hippocampal Neurons 311
R. L. Spencer and S. T. Bland

Hippocampus, Corticosteroid Effects on 321
M. Joëls and H. Karst

Hippocampus, Overview 327
M. J. Meaney and S. J. Lupien

Hiroshima Bombing, Stress Effects of 332
R. J. Lifton

HIV Infection/AIDS 336
B. W. Dixon

Holocaust Survivors, Experiences of 339
P. Valent

Holocaust, Stress Effects of 342
P. Valent

Homeostasis 347
B. S. McEwen

Homosexuality, Stress and 348
J. Drescher

Hostility 354
L. H. Powell and K. Williams

HPA Alterations in PTSD 359
R. Yehuda

Hurricane Katrina Disaster, Stress Effects of 364
C. Piotrowski

11β-Hydroxysteroid Dehydrogenases 368
J. R. Seckl

Hyperreactivity (Cardiovascular) 372
A. Georgiades

Hypertension 376
A. Steptoe

Hyperthermia 381
J. Roth

Hyperthyroidism 388
I. M. Lesser and D. L. Flores

Hyperventilation 390
C. Bass

Hypnosis 394
W. G. Whitehouse, E. C. Orne and M. T. Orne

Hypocortisolism and Stress 400
C. M. Heim and U. M. Nater

Hypoglycemia 408
B. M. Frier

Hypotension, Hypovolemia, and Septic Shock 413
A. Beishuizen, A. B. Johan Groeneveld and I. Vermes

Hypothalamic-Pituitary-Adrenal Axis 421
M. F. Dallman, S. Bhatnagar and V. Viau

Hypothermia 428
M. J. Taylor

Hypothyroidism 439
R. T. Joffe

Hysteria 442
K. Pajer

I

Immobilization Stress 445
R. Kvetnansky and R. McCarty

Immune Cell Distribution, Effects of Stress on 449
F. S. Dhabhar

Immune Function, Stress-Induced Enhancement 455
F. S. Dhabhar

Immune Response 462
P. J. Delves

Immune Suppression 470
P. Prolo and F. Chiappelli

Immune Surveillance – Cancer, Effects of Stress on 477
D. Spiegel and F. S. Dhabhar

Immune System, Aging 481
M. A. Horan

Immunity 485
F. Chiappelli

Impact of Terrorism on the Development of Mental Health Symptoms 493
R. Yehuda

Impotence, Stress and 497
M. R. Gignac, G. M. Rooker and J. K. Cohen

Impulse Control 500
I.-M. Blackburn

Incest 502
A. P. Mannarino

Income Levels and Stress 506
S. V. Subramanian and I. Kawachi

Indigenous Societies 511
W. W. Dressler

Industrialized Societies 517
J. Siegrist

Infection 521
H. Friedman and S. H. Pross

Inflammation 530
G. Z. Feuerstein, R. R. Ruffolo, C. Coughlin, J. Wang and D. Miller

Instinct Theory 535
R. Gardner Jr.

Integrative Medicine (Complementary and Alternative Medicine) 537
D. Spiegel

Interactions Between Stress and Drugs of Abuse 540
P. V. Piazza and M. Le Moal

J

Job Insecurity: The Health Effects of a Psychosocial Work Stressor 549
J. E. Ferrie and P. Martikainen

K

Kidney Function 557
W. J. Welch

Korean Conflict, Stress Effects of 563
C. A. Goguen and M. J. Friedman

L

Learned Helplessness 567
D. M. Isaacowitz and M. E. P. Seligman

Learning and Memory, Effects of Stress on 571
M. Lindau, O. Almkvist and A. H. Mohammed

Left Ventricular Mass 577
T. G. Pickering

***Leishmania*, Stress Response in** 579
M. Shapira

Leptin, Adiponectin, Resistin, Ghrelin 584
A. Gavrila, D. Barb and C. S. Mantzoros

Leukocyte Trafficking and Stress 592
S. Hong, M. U. Goebel and P. J. Mills

Life Events and Health 599
P. Surtees and N. Wainwright

Life Events Scale 603
E. Wethington

Lockerbie Air Crash, Stress Effects of 608
H. Livingston and M. Livingston

Loss Trauma 612
A. Bifulco

Lymph Nodes 616
T. L. Whiteside

Lymphocytes 622
N. R. Rose

M

Macrophage Antimycobacterial Activity, Effects of Stress on 629
C. S. Boomershine and B. S. Zwilling

Macrophages 634
W. P. Fehder, F. Tuluc, W.-Z. Ho and S. D. Douglas

Major Depressive Disorder 640
A. B. Negrão and P. W. Gold

Male Partner Violence 645
M. Ingram, N. P. Yuan and M. P. Koss

Marital Conflict 651
P. T. McFarland and A. Christensen

Marital Status and Health Problems 653
I. M. A. Joung

Marriage 660
A. C. Yoneda and J. Davila

Maternal Deprivation 667
R. L. Huot, C. O. Ladd and P. M. Plotsky

Medical Profession and Stress 674
K. G. Power and V. Swanson

Meditation and Stress 678
J. L. Kristeller

Membrane Glucocorticoid Receptors 686
P. J. Gasser and M. Orchinik

Memory and Stress 693
S. J. Lupien and F. S. Maheu

Memory Impairment 699
A. J. Parkin†

Menopause and Stress 703
N. E. Avis

Menstrual Cycles and Stress 706
R. Suri and L. Altshuler

Mental Stress Testing 712
P. G. Saab, K. A. Kline and J. R. McCalla

Metabolic Syndrome 717
L. Keltikangas-Järvinen

Metabolic Syndrome and Stress 721
R. Rosmond

Metals, Oxidative Stress, and Brain Biology 724
D. I. Finkelstein, T. Lynch, S. Wilkins, R. A. Cherny and A. I. Bush

Metastasization 726
M. K. Demetrikopoulos

Metyrapone: Basic and Clinical Studies 730
E. A. Young

Migraine 733
N. M. Ramadan

Mineralocorticoid Receptor Polymorphisms 744
R. H. DeRijk and E. R. de Kloet

Minorities and Stress 748
I. Mino, W. E. Profit and C. M. Pierce

Mitochondria 754
I. Manoli, S. Alesci and G. P. Chrousos

Monoamine Oxidase 761
P. Huezo-Diaz and I. W. Craig

Motor Vehicle Accidents, Stress Effects of 764
T. C. Buckley and E. B. Blanchard

Mucosal Secretory Immunity, Stress and 768
J. A. Bosch and D. Carroll

Multi Drug Resistance P Glycoprotein and other Transporters 774
E. C. M. de Lange

Multiple Personality Disorder 783
R. P. Kluft

Multiple Sclerosis 790
A. T. Reder

Multiple Trauma 795
P. N. Soucacos and E. O. Johnson

Musculoskeletal Problems and Stress 800
S. Svebak

Myopathy 807
G. A. Small

N

Natural Killer (NK) Cells 815
T. L. Whiteside, M. Boyiadzis and R. B. Herberman

Negative Affect 822
A. A. Stone and A. A. Gorin

Neighborhood Stress and Health 825
A. V. Diez Roux

Nelson's Syndrome 828
A. Stathopoulou, K. Dimitriou and G. Kaltsas

Neural Stem Cells 832
U. S. Sohur, J. G. Emsley, B. D. Mitchell and J. D. Macklis

Neurodegenerative Disorders 840
M. F. Mendez and A. M. McMurtray

Neurodevelopmental Disorders in Children 844
N. J. Rinehart and B. J. Tonge

Neuroendocrine Systems 851
G. Fink

Neurogenesis 865
P. Tanapat and E. Gould

†Deceased.

Neuroimaging and Emotion 870
N. A. Harrison and H. D. Critchley

Neuroimmunomodulation 879
N. R. Rose

Neuroinflammation 889
G. Z. Feuerstein, R. R. Ruffolo and L. J. Rutkowski

Neuropeptide Y 895
C. Carvajal, Y. Dumont and R. Quirion

Neuropeptides, Stress-Related 903
A. J. Tilbrook

Neurosis 908
D. J. Lynn

Neuroticism, Genetic Mapping of 910
M. W. Nash

Neuroticism Response to Stress, Genetic Mapping of Mice 914
M. R. Munafò and J. Flint

Night Shiftwork 917
T. Åkerstedt and G. Lindbeck

Nightmares 920
M. Hirshkowitz and A. Sharafkhaneh

Nitric Oxide 925
S. M. McCann

Northern Ireland, Post Traumatic Stress Disorder in 933
P. Bell

Nuclear Warfare, Threat of 937
J. Thompson

Nutrition 940
J. E. Morley

VOLUME 3

O

Obesity, Stress and 1
R. Rosmond

Obsessive–Compulsive Disorder 3
R. T. Rubin and B. J. Carroll

Oklahoma City Bombing, Stress Effects of 8
C. Piotrowski and S. J. Vodanovich

Opioids 11
J. A. Russell

Optimism, Pessimism, and Stress 26
M. F. Scheier and C. S. Carver

Orexin 29
W. K. Samson, M. M. White and A. V. Ferguson

Organ Transplantation, Stress of 35
M. A. Dew, A. F. DiMartini and R. L. Kormos

Oxidative Stress 45
H. Sies and D. Jones

Oxidative Stress and Acidosis, Molecular Responses to 49
C. Mobbs

Oxidative Stress and Aging 52
C. Mobbs

Oxytocin 54
G. Leng and N. Sabatier

P

Pain 63
H. J. Strausbaugh and J. D. Levine

Panic Disorder and Agoraphobia 66
J. C. Ballenger

Paranoia 70
P. Kinderman

Paraventricular Nucleus 75
L. W. Swanson

Parenting, Stress of 79
K. D. Jennings and Laura J. Dietz

Parkinson's Disease 84
S. Baser

Peacekeeping 88
B. Litz and S. Maguen

Peptides 91
E. P. Zorrilla and G. F. Koob

Perinatal Dexamethasone 97
E. Fuchs

Persian Gulf War, Stress Effects of 102
S. M. Southwick, D. Vojvoda, C. A. Morgan III and D. Lipschitz

Personality Processes 108
R. J. Larsen

Pharmacological Treatments of Stress 113
O. G. Cameron

Pheromones 119
A. Kumar, C. A. Dudley, S. Chakravarty and R. L. Moss

Pituitary Regulation, Role of 127
F. A. Antoni

Police, Stress in 131
R. J. Burke

Posttraumatic Stress Disorder – Clinical 135
N. C. Feeny, L. R. Stines and E. B. Foa

Posttraumatic Stress Disorder – Neurobiological basis for 140
M. Barad

Posttraumatic Stress Disorder in Children 145
A. M. La Greca

Posttraumatic Stress Disorder, Delayed 150
A. Holen

Posttraumatic Stress Disorder, Neurobiology of 152
J. D. Bremner

Posttraumatic Therapy 157
A. M. Rasmusson, C. M. Monson and P. A. Resick

Pregnancy – Maternal and Perinatal Stress – Effects of 165
P. Smirnaki and M.-A. Magiakou

Premenstrual Dysphoric Disorder 173
S. Nowakowski, P. Haynes and B. L. Parry

Pre-pulse Inhibition 180
M. van den Buuse

Pressure, Effects of Extreme High and Low 184
R. J. Værnes

Primate Hierarchies and Personality 194
R. M. Sapolsky

Primate Models, Behavioral–Immunological Interactions 199
M. L. Laudenslager and S. Tiefenbacher

Primate Models, Cardiovascular Disease 204
J. R. Kaplan

Primate Models, Overview 211
D. M. Lyons

Primates: Rearing and Effects of Stress on Primate CNS Function 214
T. K. Newman and C. S. Barr

Prison 217
D. L. Whitehead and A. Steptoe

Prisoners of War 223
C. Tennant

Problem-Solving Skills Training 227
A. M. Nezu, C. M. Nezu and T. J. D'Zurilla

Prolactin and Stress 231
G. Tolis, G. Rombopoulos, D. Kaltsas, E. Katounda, V. Kaltzidou and N. Angelopoulos

Pro-opiomelanocortin (POMC) 233
A. B. Bicknell

Prostaglandins 240
S. Moshonov, U. Zor and Z. Naor

Proteases in Prokaryotes and Eukaryotic Cell Organelles 247
E. Conway de Macario and A. J. L. Macario

Proteases in the Eukaryotic Cell Cytosol 252
E. Conway de Macario and A. J. L. Macario

Protein Synthesis 258
M. A. Brostrom and C. O. Brostrom

Proteosome 266
M. Maldonado and J. Wang

Psoriasis 271
A. B-. Kirschbaum

Psychoanalysis 274
P. Roazen

Psychological Stressors, Overview 278
S. M. Monroe and G. M. Slavich

Psychoneuroimmunology 284
R. Dantzer

Psychosocial Factors and Stress 288
J. Siegrist

Psychosomatic Heart Disease: Role of Sympathetic and Sympathoadrenal Processes 292
G. W. Lambert and T. Dawood

Psychosomatic Medicine 296
T. Theorell

Psychotherapy 302
G. C. Smith

Psychotic Disorders 308
J. Ventura

Q

Quality of Life 317
S. M. Skevington

R

Racial Harassment/Discrimination 321
D. R. Williams and S. A. Mohammed

Recovery from Stress 327
S. Sonnentag and C. Fritz

Reductive Stress 331
J. P. Kehrer

Reenactment Techniques 336
N. Wong

Refugees, Stress in 338
J. D. Kinzie

Regional Blood Flow, Stress Effects 341
W. W. Blessing

Relaxation Techniques 345
W. G. Whitehouse, E. C. Orne and M. T. Orne

Religion and Stress 351
S. Packer

9/11, Religion and Stress 357
S. Packer

Remodelling of Neuronal Networks by Stress 364
E. Fuchs and G. Flügge

Renal and Adrenocortical Actions of Dopamine 371
B. C. Williams, Y.-C. Lo and S. W. Walker

Reproduction, Effects of Social Stress On 374
C. A. Shively

Reproductive Dysfunction in Primates, Behaviorally Induced 380
J. L. Cameron

Resistance 386
L. M. Zabarenko

Restraint Stress 389
R. J. Servatius, G. Salameh, K. M. Coyle and W. P. Paré

Revenge Fantasies 391
M. Horowitz and S. Meffert

Rheumatic Disorders 395
A. T. Masi and J. C. Aldag

S

Salivary Cortisol 405
C. Kirschbaum and D. H. Hellhammer

Salt Appetite 409
R. S. Weisinger and N. Chen

Schizophrenia 418
M. van den Buuse and D. Copolov

School Stress and School Refusal Behavior 422
C. A. Kearney, L. C. Cook and G. Chapman

School Violence and Bullying 425
J. Juvonen

Seasonal Changes in Stress Responses 427
R. J. Nelson and L. B. Martin II

Seasonal Rhythms 432
L. M. Romero

Secretagogue 435
G. Fink

Selective Serotonin Reuptake Inhibitors (SSRIs) 440
M. Gitlin

Self-Esteem, Stress, and Emotion 443
D. Roger

Selye, Hans 448
B. Tuchweber and P. Bois

Sepsis, Acute Respiratory Distress Syndrome, and Glucocorticoid Resistance 450
G. U. Meduri

Serotonin 457
B. C. Williams and Y.-C. Lo

Serotonin in Stress 461
S. Kusljic and M. van den Buuse

Serotonin Transporter Genetic Modifications 465
K. Sugden

Sex Differences in Human Stress Response 469
B. M. Kudielka, D. H. Hellhammer and C. Kirschbaum

Sex Steroids, Response to Stress and Susceptibility to Depression 474
M. V. Seeman and L. E. Ross

Sex-Specific Effects of Early Social Stress in Mammals: A Study in Guinea Pigs 479
S. Kaiser and N. Sachser

Sexual Assault 484
N. C. Feeny, T. J. Linares and E. B. Foa

Sexual Dysfunction 490
W. T. O'Donohue and L. Woodward Tolle

Sexual Offenders 494
F. M. Saleh and H. M. Malin

Sickle Cell Disease and Stress 502
K. Midence

Sleep Loss, Jet Lag, and Shift Work 504
E. Van Cauter

Sleep, Sleep Disorders, and Stress 506
A. N. Vgontzas, S. Pejovic and M. Karataraki

Smoking and Stress 515
F. J. McClernon and D. G. Gilbert

Social Capital 520
I. Kawachi

Social Networks and Social Isolation 523
L. F. Berkman

Social Status and Stress 528
D. de Ridder

Social Stress, Animal Models of 533
J. Bugajski, A. Gądek-Michalska and A. J. Bugajski

Social Support 539
T. C. Antonucci, J. E. Lansford and K. J. Ajrouch

Social Support in Trauma 542
K. O'Donnell and A. Steptoe

Somatic Disorders 545
F. Creed

Space, Health Risks of 548
V. A. Convertino

Spinal Cord Injury, Physical Stress of 554
S. Bhatia and M. Quigley

Spinal Cord Injury, Psychological Stress of 557
S. D. Schnakenberg-Ott and M. R. Lovell

Startle Response 561
C. O. Ladd, P. M. Plotsky and M. Davis

Steroid Hormone Receptors 568
M. Beato and J. Klug

Steroid Hydroxylases 581
F. H. de Jong

Strain Differences in Stress Response in Rodents 585
P. Mormède

Stress and Anxiety: Treatment with 2nd Generation Antipsychotic Drugs 587
J. S. Ballon, J. A. Boyd and D. A. Wirshing

Stress and CNS Arousal: Genomic Contributions 591
A. C. Ribeiro and D. W. Pfaff

Stress Effect of Assisted Reproduction 596
F. M. Helmerhorst, R. Sibug and E. R. de Kloet

Stress Effects, Overview 599
A. Steptoe

Stress Generation 601
J. E. Roberts and J. A. Ciesla

Stress Hyporesponsive Period 606
S. Levine, E. R. de Kloet, G. Dent and M. S. Schmidt

Stress in University Students 612
B. Andrews and J. Hejdenberg

Stress Induced Anovulation 615
S. L. Berga and T. L. Loucks

Stress Management and Cardiovascular Disease 631
D. Lane and D. Carroll

Stress Management, CAM Approach 636
K. Krebs

Stress of Self Esteem 640
J. C. Pruessner, S. Wuethrich and M. W. Baldwin

Stress System Balance Hypothesis 646
E. R. de Kloet

Stress, Beneficial Effects of 650
S. Joseph and P. A. Linley

Stress, Definitions and Concepts of 653
B. S. McEwen

Stress, Insulin Resistance, and Type II Diabetes 654
C. Tsigos and I. Kyrou

Stress, NPY, and Cardiovascular Diseases 660
Z. Zukowska

Suicide Terrorism, Genesis of 667
A. Speckhard

Suicide, Biology of 677
M. A. Oquendo, L. Giner and J. J. Mann

Suicide, Psychology of 684
D. Lester and R. L. Walker

Suicide, Sociology of 689
D. Lester and R. M. Fernquist

Surgery and Stress 693
C. Vögele

Survivor Guilt 695
P. Valent

Sympathetic Nervous System 697
D. S. Goldstein

Synthetic Glucocorticoids 704
A. M. Karssen and E. R. de Kloet

Systemic Lupus Erythematosus 708
D. J. Wallace

T

Teaching and Stress 713
E. R. Greenglass

Temperature Effects 717
J. Roth

Terrorism 719
P. Bell

Thermal Stress 723
C. M. Blatteis

Thermotolerance, Thermoresistance, and Thermosensitivity 726
S. Gatti McArthur, D. Alberati and T. Bartfai

Three Mile Island, Stress Effects of 735
A. L. Dougall and A. Baum

Thymus 738
M. S. Vacchio

Thyroid Hormones 743
T. J. Visser and E. Fliers

Torture 749
I. Genefke, H. Marcussen and O. V. Rasmussen

Transport-Related Stress 756
R. G. Smart

Trans-sexualism 760
R. A. Allison

Trauma and Memory 765
B. A. van der Kolk

Trauma Group Therapy 767
J. Dwyer and L. G. Martin

Traumatic Stress and Posttraumatic Stress Disorder, the Israeli Experience 771
E. Klein and J. Zohar

Trier Social Stress Test 776
B. M. Kudielka, S. Wüst, C. Kirschbaum and D. H. Hellhammer

Type A Personality, Type B Personality 782
W. S. Shaw and J. E. Dimsdale

U

Ulceration, Gastric 787
R. Murison and A. M. Milde

Ultradian Rhythms 791
S. L. Lightman

Understimulation/Boredom 794
V. J. Sutherland

Unemployment, Stress, and Health 797
M. Bartley

Urocortins 804
É. M. Fekete and E. P. Zorrilla

V

Vaccination 813
V. E. Burns, A. C. Phillips and K. M. Edwards

Vasoactive Peptides 817
W. K. Samson and M. M. White

Vasopressin 824
L. P. Renaud

Vietnam Veterans, Postwar Experiences and Health Outcomes 830
J. A. Boscarino

Violence 838
E. K. Englander

Viral Virulence and Stress 842
A. J. L. Macario and E. Conway de Macario

Viruses and Stress 850
G. Z. Feuerstein, R. R. Ruffolo and B.-N. David

W

Waist–Hip Ratio 853
R. Rosmond

War Stress in the Former Yugoslavia 855
M. Flögel, S. Šupraha Goreta and G. Lauc

War, Suicide and Sacrifice 860
V. Hazboun

War-Related Posttraumatic Stress Disorder, Treatment of 865
L. B. Slone and M. J. Friedman

Work–Family Balance 868
J. G. Grzywacz and A. B. Butler

Workplace Stress 871
U. Lundberg

VOLUME 4

Contributors *1*
Subject Index *21*

CONTENTS BY SUBJECT AREA

ANIMAL STUDIES AND MODELS

Animal Models (Nonprimate) for Human Stress
Circadian Rhythms, Effects of Prenatal Stress in Rodents
Fish, Stress in
Immobilization Stress
Pre-pulse inhibition
Primate Hierarchies and Personality
Primate Models, Behavioral-Immunological Interactions
Primate Models, Cardiovascular Disease
Primate Models, Overview
Primates: Rearing and Effects of Stress on Primate CNS Function
Sex-Specific Effects of Early Social Stress in Mammals: a Study in Guinea Pigs

CONFLICT, WAR, TERRORISM

Chemical Warfare
Gulf War Syndrome, Psychological and Chemical Stressors
Hiroshima Bombing, Stress Effects of
Holocaust, Stress Effects of
Impact of Terrorism on the Development of Mental Health Symptoms
Korean Conflict, Stress Effects of
Lockerbie Air Crash, Stress Effects of
Northern Ireland, Post Traumatic Stress Disorder in
Nuclear Warfare, Threat of
Oklahoma City Bombing, Stress Effects of
Persian Gulf War, Stress Effects of
Prisoners of War
9/11, Religion and Stress
Suicide Terrorism, Genesis of
Terrorism
Torture
Traumatic Stress and Posttraumatic Stress Disorder; the Israeli Experience
Vietnam Veterans, Postwar Experiences and Health Outcomes
War Stress in the Former Yugoslavia
War, Suicide and Sacrifice

DISASTERS

Airline Accidents
Chernobyl, Stress Effects of
Disasters and Mass Violence, Public, Effects of
Disaster Syndrome
Earthquakes, Stress Effects of
Floods, Stress Effects of
Hurricane Katrina Disaster, Stress Effects of
Motor Vehicle Accidents, Stress Effects of
Three Mile Island, Stress Effects of

DIURNAL, SEASONAL AND ULTRADIAN RHYTHMS

Circadian rhythm effects on cardiovascular and other stress-related events
Circadian Rhythms, Genetics of
Night Shiftwork
Seasonal Changes in Stress Responses
Seasonal Rhythms
Sleep Loss, Jet Lag, and Shift Work
Sleep, Sleep Disorders, and Stress
Ultradian Rhythms

DRUGS (EFFECTS)

Alcohol, Alcoholism and Stress: A Psychobiological Perspective
Drug Use and Abuse
Interactions between Stress and Drugs of Abuse
Opioids

DRUGS (TREATMENT)

Antipsychotic Drugs and Stress
Anxiolytics
Benzodiazepines
Beta-Adrenergic Blockers
Pharmacological Treatments of Stress
Selective Serotonin Reuptake Inhibitors (SSRIs)
Stress and Anxiety: treatment with 2nd generation antipsychotic drugs

GENERAL CONCEPTS AND MODELS

Alarm Phase and General Adaptation Syndrome
Allostasis and Allostatic Load
Behavior, Overview
Conservation of Resources Theory
Control and Stress
Demand–Control Model
Effort–Reward Imbalance Model
Environmental Factors
Evolutionary Origins and Functions of the Stress Response
Fight-or-Flight Response
Instinct Theory
Life Events Scale
Psychological Stressors, Overview
Remodelling of neuronal networks by stress
Selye, Hans
Stress, Definitions and Concepts of
Stress Effects, Overview
Stress System Balance Hypothesis
Stress, Beneficial Effects of

GENETICS AND GENOMICS

Circadian clock genes as modulators of sensitivity to genotoxic stress
Corticosteroid Receptor Genes: Functional Dissection in Mice
Corticotropin Releasing Factor Receptor Deficiency in Mice
Drosophila Genes and Anoxia
Expression Profiling of Stress Responsive Gene Patterns
Gene Environment Interactions in Early Development
Genetic Factors and Stress
Genetic Polymorphisms in Stress Response
Genetic Predispositions to Stressful Conditions
Genetic Testing and Stress
Genetic variation of HPA axis activity and function in farm animals
Glucocorticoid Receptor Mutant Mice as Models for Stress-Induced Affective Disorders
Glucocorticoid Receptor Mutations and Polymorphisms
Mineralocorticoid Receptor Polymorphisms
Neuroticism Genetic Mapping of
Neuroticism Response to Stress, Genetic Mapping of Mice
Serotonin Transporter Genetic Modifications
Strain Differences in Stress Response in Rodents
Stress and CNS Arousal; Genomic Contributions

HUMAN COGNITION, EMOTION AND BEHAVIOR

Adolescence
Aggression
Aggressive Behavior
Aging and Psychological Stress
Anger
Anxiety
Bereavement
Burnout
Captivity, Adaptation to
Captivity, Recovery from
Caregivers, Stress and
Childbirth and Stress
Cognition and Stress
Combat, Acute Reactions to
Combat Reaction, Chronic
Combat Stress Reaction
Concentration Camp Survivors
Coping Skills
Coping and Stress: a Lens and Filter Model
Cortisol Awakening Response
Death Anxiety
Dental Stress
Depersonalization: Systematic Assessment
Diet and Stress, Non-Psychiatric
Dissociation
Distress
Emergency Personnel, Stress in
Emotional Inhibition
Emotions: Structure and Adaptive Functions
Fatigue and Stress
Fear
Fear and the Amygdala
Firefighters, Stress in
Grieving
Holocaust Survivors, Experiences of
Homosexuality, Stress and
Hostility
Impotence, Stress and
Impulse Control
Industrialized Societies
Learning and Memory, Effects of Stress on
Marriage

Medical Profession and Stress
Memory Impairment
Menopause and Stress
Neuroimaging and Emotion
Nightmares
Optimism, Pessimism and Stress
Pain
Parenting, Stress of
Peacekeeping
Police, Stress in
Prison
Psychosomatic Medicine
Recovery from Stress
Refugees, Stress in
Religion and Stress
Self-Esteem, Stress, and Emotion
Somatic Disorders
Stress Generation
Stress in University Students
Suicide, Psychology of
Suicide, Sociology of
Surgery and Stress
Survivor Guilt
Teaching and Stress
Trauma and Memory
Type A Personality, Type B Personality
Understimulation/Boredom

HUMAN HEALTH AND PHYSICAL ILLNESS

Amnesia
Adrenal Insufficiency
AIDS
Alzheimer's Disease
Arthritis
Asthma
Atherosclerosis
Attention-Deficit/Hyperactivity Disorder, Stress and
Breast Cancer
Cancer
Cardiovascular Disease, stress and
Chaperonopathies
Chronic Fatigue Syndrome
Common Cold and Stress
C-Reactive Protein
Cytokines, Chronic Stress, and Fatigue
Cushing's Syndrome, Medical Aspects
Cushing's Syndrome, Neuropsychiatric Aspects
Dermatological Conditions
DHEA
Diabetes, Type I
Disease, Stress Induced
Eclampsia and preeclampsia
Endometriosis
Enuresis
Epilepsy
Fibromyalgia
Graves' Disease (Thyrotoxicosis)
Health Behavior and Stress
Heart Disease/Attack
Heart Failure, Stress Effects
Hemostasis and Stress
Hyperreactivity (Cardiovascular)
Hypertension
Hyperthermia
Hyperthyroidism
Hyperventilation
Hypoglycemia
Hypotension, Hypovolemia, and Septic Shock
Hypothermia
Hypothyroidism
Hysteria
Metastasization
Migraine
Multiple Trauma
Nelson's Syndrome
Neurodegenerative Disorders
Organ Transplantation, Stress of
Parkinson's Disease
Perinatal Dexamethasone
Pregnancy, Maternal and Perinatal Stress, Effects of
Premenstrual Dysphoric Syndrome
Prolactin and stress
Psychosomatic heart disease: role of sympathetic
and sympathoadrenal processes
Rheumatic Disorders
Sickle Cell Disease and Stress
Space, Health Risks of
Spinal Cord Injury, Physical Stress of
Spinal Cord Injury, Psychological Stress of
Stress Effect of Assisted Reproduction
Stress Induced Anovulation
Stress, Insulin Resistance and Type II Diabetes
Stress, NPY and Cardiovascular Diseases
Ulceration, Gastric

HUMAN MENTAL HEALTH AND PSYCHOPATHOLOGY

Acute Stress Disorder and Posttraumatic Stress
Disorder
Adjustment Disorders
Adolescent suicide
Affective Disorders
Antisocial Disorders
Anxiety
Arthritis – Psychological
Borderline Personality Disorder
Child Abuse
Child Sexual Abuse

Comorbid Disorders and Stress
Corticotropin-Releasing Factor Circuitry in the Brain – Relevance for Affective Disorders and Anxiety
Defensive Behaviors
Depression and Coronary Heart Disease
Depression and Manic-Depressive illness
Depression and Stress, Role of n-3 and n-6 Fatty Acids
Depression, Immunological Aspects
Depression Models
Diet and Stress, Psychiatric
Disaster Syndrome
Dissociation
Eating Disorders and Stress
Elder Abuse
Ethnicity, Mental Health
HPA alterations in PTSD
Hypocortisolism and Stress
Incest
Learned Helplessness
Life Events and Health
Loss Trauma
Male Partner Violence
Major Depressive Disorder
Multiple Personality Disorder
Negative Affect
Neurodevelopmental Disorders in Children
Neurosis
Obsessive–Compulsive Disorder
Panic Disorder and Agoraphobia
Paranoia
Posttraumatic Stress Disorder – Clinical
Posttraumatic Stress Disorder, Delayed
Posttraumatic Stress Disorder in Children
Posttraumatic Stress Disorder – neurobiological basis for
Posttraumatic Stress Disorder, Neurobiology of
Psychotic Disorders
Revenge Fantasies
Schizophrenia
Sexual Assault
Sexual Offenders
Social support in trauma
Stress of Self Esteem
Suicide, Biology of
Trans-sexualism
Violence

IMMUNOLOGY, INFECTION AND INFLAMMATION

Antibody Response
Autoimmunity
Corticotropin-releasing factor (CRF) family of neuropeptides – role in inflammation
Cytokines
Cytokines, Stress, and Depression
Cytotoxic Lymphocytes
Herpesviruses
HIV Infection/AlDS
Immune Cell Distribution, Effects of Stress on
Immune Function, Stress-lnduced Enhancement of
Immune Response
Immune Suppression
Immune Surveillance – Cancer, Effects of Stress on
Immune System, Aging
Immunity
Infection
Inflammation
Leishmania, Stress Response in
Leukocyte Trafficking and Stress
Lymphocytes
Mucosal Secretory Immunity, Stress and
Multiple Sclerosis
Natural Killer (NK) Cells
Neuroimmunomodulation
Neuroinflammation
Psoriasis
Psychoneuroimmunology
Sepsis, Acute Respiratory Distress Syndrome and Glucocorticoid Resistance
Systemic Lupus Erythematosus
Vaccination
Viral Virulence and Stress
Viruses and Stress

LABORATORY STUDIES AND TESTS

Antimineralocorticoid Challenge
Dex-CRH Test
Dexamethasone Suppression Test (DST)
Mental Stress Testing
Metyrapone: Basic and Clinical Studies
Trier Social Stress Test
Waist–Hip Ratio

THERAPIES

Aerobics, Exercise and Stress Reduction
Behavior Therapy
Cancer Treatment
Desensitization
Exercise
Family Therapy
Group Therapy
Hypnosis
Indegrative Medicine (Complementary and Alternative Medicine)

Meditation and Stress
Problem-Solving Skills Training
Psychosomatic Medicine
Reenactment Techniques
Relaxation Techniques
Stress Management and Cardiovascular Disease
Stress Management, CAM Approach
Trauma Group Therapy

PHYSIOLOGICAL, PHARMACOLOGICAL AND BIOCHEMICAL ASPECTS

Acute Stress Response: Experimental
Acute Trauma Response
Adenylyl Cyclases and Stress Response
Adrenal Cortex
Adrenaline
Adrenal Medulla
Adrenocorticotropic Hormone (ACTH)
Adrenocortical Function, factors controlling development thereof
Aging and Stress, Biology of
Aging and Adrenocortical Factors
Alcohol and Stress: Social and Psychological Aspects
Aldosterone and Mineralocorticoid Receptors
Ambulatory Blood Pressure Monitoring
Amenorrhea
Amygdala
Anatomy of the HPA Axis
Androgen Action
Angiotensin
Angiotensin receptors
Annexin
Anti-CRF
Antidepressant Actions on Glucocorticoid Receptors
Apoptosis
Arterial Baroreflex
Autonomic Nervous System
Autotolerance
Avoidance
Beta-Endorphin
Blood–Brain Barrier, Stress and
Blood Pressure
Brain and Brain Regions
Brain Natriuretic Peptide (BNP)
Brain Trauma
Calbindin
Calcium-Dependent Neurotoxicity
Calcium, Role of
Cardiovascular System and Stress
Catecholamines
Central Stress Neurocircuits
Cerebral Metabolism, Brain Imaging
Chaperone Proteins and Chaperonopathies
Cholesterol and Lipoproteins
Comparative Anatomy and Physiology
Congenital Adrenal Hyperplasia (CAH)
Corticosteroid-Binding Globulin (Transcortin)
Corticosteroid Receptors
Corticosteroids and Stress
Corticotropin Releasing Factor (CRF)
Corticotropin Releasing Factor-Binding Protein
Corticotropin Releasing Factor Receptors
Critical Thermal Limits
Dopamine, Central
Drosophila Studies
Electrodermal Activity
Endocrine Systems
Estrogen
Ethanol and Endogenous Opioids
Excitatory Amino Acids
Excitotoxins
Familial Patterns of Stress
Febrile Response
Feedback Systems
Fibrinogen and Clotting Factors
Feeding Circuitry (and Neurochemistry)
Fetal Stress
Food Intake and Stress, Human
Food Intake and Stress, Non-Human
Food Shift Effect
GABA (Gamma Aminobutyric Acid)
Gastrointestinal Effects
Gender and Stress
Ghrelin and Stress Protection
Glia or Neuroglia
Glucocorticoid Negative Feedback
Glucocorticoid Effects on Memory: the Positive and Negative
Glucocorticoids – Adverse Effects on the Nervous System
Glucocorticoids, Effects of Stress on
Glucocorticoids, Overview
Glucocorticoids, Role in Stress
Glucose Transport
Glycobiology of Stress
Gonadotropin Secretion, Effects of Stress on
Heart Rate
Heat Resistance
Heat Shock Genes, Human
Heat Shock Proteins: HSP60 Family Genes
Heat Shock Response, Overview
Hippocampal Neurons
Hippocampus, Corticosteroid Effects on
Hippocampus, Overview
Homeostasis
11β-Hydroxysteroid Dehydrogenases
Hypothalamic-Pituitary-Adrenal Axis

Instinct Theory
Kidney Function
Left Ventricular Mass
Leptin, Adiponectin, Resistin, Ghrelin
Lymph Nodes
Macrophage Antimycobacterial Activity, Effects of Stress on
Macrophages
Maternal Deprivation
Membrane Glucocorticoid Receptors
Memory and Stress
Menstrual Cycles and Stress
Metabolic Syndrome
Metals, Oxidative Stress and Brain Biology
Mitochondria
Monoamine Oxidase
Multi Drug Resistance P Glycoprotein and other Transporters
Musculoskeletal Problems and Stress
Myopathy
Neural Stem Cells
Neuroendocrine Systems
Neurogenesis
Neuropeptide Y
Neuropeptides, Stress-Related
Nitric Oxide
Nutrition
Obesity, Stress and
Orexin
Oxidative Stress
Oxidative Stress and Acidosis, Molecular Responses to
Oxidative Stress and Aging
Oxytocin
Paraventricular Nucleus
Peptides
Personality Processes
Pheromones
Pituitary Regulation, Role of
Pressure, Effects of Extreme High and Low
Pro-opiomelanocortin (POMC)
Prostaglandins
Proteases in Prokaryotes and Eukaryotic Cell Organelles
Proteases in the Eukaryotic Cell Cytosol
Protein Synthesis
Proteosome
Reductive Stress
Regional Blood Flow, Stress Effects
Renal and Adrenocortical Effects of Dopamine
Reproduction, Effects of Social Stress on
Reproductive Dysfunction in Primates, Behaviorally Induced
Resistance
Restraint Stress
Salivary Cortisol
Salt Appetite
Secretagogue
Serotonin
Serotonin in Stress
Sex Differences in Human Stress Response
Sex Steroids, Response to Stress and Susceptibility to Depression
Sexual Dysfunction
Smoking and Stress
Startle Response
Steroid Hormone Receptors
Steroid Hydroxylases
Stress Hyporesponsive Period
Sympathetic Nervous System
Synthetic Glucocorticoids
Temperature Effects
Thermal Stress
Thermotolerance, Thermoresistance, and Thermosensitivity
Thymus
Thyroid Hormones
Urocortin
Vasoactive Peptides
Vasopressin

PSYCHOLOGICAL THERAPY

Cognitive Behavioral Therapy
Freud, Sigmund
Posttraumatic Therapy
Psychoanalysis
Psychotherapy
War-Related Posttraumatic Stress Disorder, Treatment of

PSYCHOSOCIAL AND SOCIOECONOMIC ASPECTS

Childhood Stress
Child Physical Abuse
Chronic Social Stress: GR Sensitivity in Leukocytes
Community Studies
Crowding Stress
Crime Victims
Crisis Intervention
Cultural Factors in Stress
Cultural Transition
Divorce, Children of
Domestic Violence
Economic Factors and Stress
Education Levels and Stress

Employee Assistance and Counseling
Environmental Stress, Effects on Human Performance
Health and Socioeconomic Status
Income Levels and Stress
Indigenous Societies
Job Insecurity; the health effects of a psychosocial work stressor
Marital Conflict
Marital Status and Health Problems
Minorities and Stress
Neighborhood Stress and Health
Psychosocial Factors and Stress
Quality of Life
Racial Harassment/Discrimination
School Violence and Bullying
School Stress and School Refusal Behavior
Social Capital
Social Networks and Social Isolation
Social Status and Stress
Social Stress, Animal Models of
Social Support
Transport-Related Stress
Unemployment, Stress and Health
Work–Family Balance
Workplace Stress

PREFACE TO FIRST EDITION

"Stress" remains one of the most frequently used but ill-defined words in the English language. Stress is a phenomenon that has quite different meanings for the politician, social scientist, physician, nurse, psychotherapist, physiologist, molecular biologist, and perhaps you and me. This diversity of meanings was one impetus for creating the *Encyclopedia of Stress,* the aim being to derive a definition of stress from a variety of expert descriptions. The second impetus was the obvious need for an up-to-date compendium on one of the most important social, medical, and psychological phenomena of our age. We were fortunate in attracting stars for our Editorial Board and a set of most distinguished contributors for the 400 or so entries—indeed, the list of contributors is a Who's Who in stress research.

We anticipate that the diversity of our readers will equal the diversity of the topics covered. They will find that the coverage of the Encyclopedia extends well beyond the general adaptation theory of Hans (Janos) Selye and the fight-or-flight response of Walter Cannon. Nevertheless, the general principles annunciated by these two great pioneers in the field still underpin our understanding of the biology of the stress phenomenon. That is, stress is a real or perceived challenge, either endogenous or exogenous, that perturbs body equilibrium or "homeostasis." The stressor may range from overcrowding, traffic congestion, violence, bereavement, redundancy, or unemployment to physical, chemical, biological, or psychological insults. Whether the person can adapt to or cope with the stress will depend on the nature and severity of the stressor and the person's physical and mental state, which in turn depends on genetic, experiential, social, and environmental factors. These issues are discussed in depth in the Encyclopedia, as are the mechanisms of coping and the impact of stress on health and predisposition to diseases such as cancer, infection, rheumatoid arthritis, heart disease, high blood pressure, and mental disorder.

Aggression remains a hallmark of human behavior, even as we move into the third millennium, and so the Encyclopedia covers several topical areas that have only recently been analyzed systematically. These include war and specific wars, posttraumatic stress disorder (PTSD; formerly thought of vaguely as "shell shock"), rape, torture, marital discord and spousal abuse, and the Holocaust. In tackling these topics we accept that our entries may not include all the nuances that are necessary for a full understanding of what these phenomena are all about and described so graphically and sensitively in Tolstoy's *War and Peace* or Pat Barker's monumental *Regeneration* trilogy on the horrific psychological traumas of the First World War. Nevertheless, an important start has been made in that we now accept that PTSD is not just a lack of "bottle" (courage or "guts"), but rather a syndrome that needs to be and can be understood within the framework of medicine and psychology.

Biologically, the stress response reflects a set of integrated cascades in the nervous, endocrine, and immune defense systems. As in most areas of biology, molecular genetics has made a significant difference in the precision with which we now understand the physiopathological processes of the stress response. And so the adage formerly applied to diabetes mellitus may now apply equally to stress: "Understand stress and you will understand medicine."

In summary, we hope that this first *Encyclopedia of Stress* will indeed define the phenomenon and at the same time provide a valuable source of information on a phenomenon that affects us all. In setting out on this adventure we were aware that there is nothing new under the sun and that stress has been around since the first biological particles, bacteria, or even viruses competed for the same mechanisms of replication.

There is a tendency for each generation to imagine that stress and its untoward effects are uniquely harsh for them, but it is not this misconception that underlies this work. Rather, the stress of "stress" itself—the massive accumulation of knowledge—made it seem propitious to bring the information together in a systematic manner that allows ready access to all who need or wish to understand the phenomenon.

The idea of producing this Encyclopedia was conceived at an Academic Press reception in San Diego held in conjunction with the annual meeting of the Society for Neuroscience in 1996. I am deeply indebted to Erika Conner for enabling conversion of the idea to a concept and then a project and to Jennifer Wrenn, Christopher Morris, and Carolan Gladden, all of Academic Press, for their enthusiasm, encouragement, and herculean efforts that converted the concept into a reality. For giving generously of their intellect, expertise, sound advice, and unstinting work, I am greatly indebted to my friends and colleagues on the Editorial Board, who made the project such a satisfying experience. To all our contributors go our profound thanks for taking time out from wall-to-wall schedules to produce their entries, which together have made this Encyclopedia.

George Fink
August 1999

PREFACE TO THE SECOND EDITION

The popularity of the first edition of the Encylopedia of Stress, the several major stressful conflicts that have occurred since 1999, and the significant advances in our knowledge of stress prompted the production of this second edition. In addition to more than 140 new articles, most of the original 400 first edition articles have been updated for the second edition to reflect, for example, advances in our understanding of corticotropin releasing factor (CRF) and urocortin receptors, ideas about the role of central CRF in the overall control of the psychological/mental as well as neuroendocrine response to stress, and the novel discoveries of the complex neuropeptide regulation of hunger and satiety. The effects of maternal and perinatal stress on subsequent body development and function in the adult, the adverse effects of stress on brain (and particularly hippocampal memory function), the role of the amygdala in fear and aggression, and the role of stress in the etiology of obesity and the metabolic syndrome all feature prominently in this new edition. 9/11 and the many new conflicts since 1999 have engendered new articles on terrorism, suicide bombers and their effects, as well as a retrospective on previous human apocalypses such as Hiroshima. Coverage of the whole field of post traumatic stress disorder (PTSD) has been expanded, reflecting the greater recognition and understanding of PTSD, and possibly related disorders such as *combat fatigue* and *burnout*. Stress and depression and anxiety, already touched on in the first edition, now feature prominently. On the conceptual side, *allostasis* (modified from Hans Selye's *heterostasis*), or stability through centrally regulated change in physiological set points, receives greater emphasis than it did in the first edition.

Overall, our understanding of stress mechanisms in the human has benefited from two major quantum leaps in technology. First, the advances in molecular genetics and genomics and sequencing of the human genome (published after the first edition had appeared) have increased the rigor and precision of our understanding of the molecular neurobiology of stress. Thus, new to the second edition are a series of articles on the genetic factors that play a role in susceptibility to stress and in various components of the stress response. Secondly, functional brain imaging has enhanced our understanding of the neurobiology of stress in the human. Many of the articles in the second edition reflect the positive impact of these two advances.

Accessibility to the second edition of the Encyclopedia of Stress will be facilitated by virtue of the fact that it is published online as well as in hard copy. My profound thanks for their selfless work go to our team of distinguished authors and to my friends and colleagues on the Editorial Board. Production of the Encyclopedia would not have been possible without our colleagues at Elsevier, and in particular Bob Donaldson and Hilary Rowe whose dedication, commitment and excellent work I am pleased to acknowledge.

George Fink
November 2006

Note on Terminology of Corticotropin Releasing Factor/Hormone and the Catecholamines

The central nervous regulation of the anterior pituitary gland is mediated by substances, mainly peptides, which are synthesized in the hypothalamus and transported to the gland by the hypophysial portal vessels. Because these compounds are transported by the blood, the term "hormone" gained acceptance in the neuroendocrine literature. The major hypothalamic peptide involved in the stress response is the

41-amino acid corticotropin releasing factor (CRF). The Endocrine Society (USA), following convention, adopted the term corticotropin releasing hormone (CRH). However, this nomenclature has been challenged. Hauger et al. (*Pharmacol Rev* 55: 21–26, 2003), in liaison with the International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification, argued that the CRF's function extends well beyond the biology of a hormone, and that therefore it should be termed corticotropin releasing factor (CRF) rather than hormone. Since the terminology of CRF versus CRH has yet to be resolved, the two terms and abbreviations are here used interchangeably, depending on author preference.

Adrenaline and *noradrenaline* are catecholamines that play a pivotal role in the stress response. These terms are synonymous with *epinephrine* and *norepinephrine*, respectively. Both sets of terms are used interchangeably in the endocrine, neuroendocrine and stress literature, and this principle has been adopted for the Encyclopedia. Style has depended on author preference, but wherever possible within-article consistency has been ensured.

George Fink
November 2006

Hauger, R. L., Grigoriadis, D. E., Dallman, M. F., Plotsky, P. M., Vale, W. W., and Dautzenberg, F. M. (2003). International Union of Pharmacology. XXXVI. Current status of the nomenclature for receptors for corticotropin-releasing factor and their ligands. *Pharmacological Reviews* **55**(1), 21–26.

FOREWORD

This is the second edition of the *Encyclopedia of Stress*, the three heavy volumes that first appeared in 2000 under the overall editorial guidance of George Fink. From its original coining in reference to rather unusual and unpleasant situations, the word *stress* is now quasi synonymous with *life*, i.e. modern life! This second edition of the *Encyclopedia of Stress* presents several hundreds of articles written by experts in that particular field, especially for this series, discussing, clarifying, explaining to the average-reader as well as to anyone more educated in one field or another, the extraordinary involvement of stress and our normal or abnormal reactions to it, both in our physical body and its accompanying mind. The table of contents is truly amazing, from the stress of being born to the stress of dying and so many circumstances in between, with clear presentations of the pertinent, current knowledge about them. This is truly an encyclopedia where, surely, you will find some explanation and answer to your own stresses (and those of others).

Roger Guillemin, MD, PhD
Nobel laureate in Physiology & Medicine

O

Obesity, Stress and

R Rosmond
Björndammsterrassen 41, Partille, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Björntorp, R Rosmond and J Uddén, volume 3, pp 75–76, © 2000, Elsevier Inc.

Stress-Eating
Stress-Eating and Obesity
Mechanism of Stress-Eating

Glossary

Hypothalamic-pituitary-adrenal axis	The hormonostatic endocrine axis that regulates cortisol production and secretion.
Leptin	An adipocyte-produced hormone that is involved in the regulation of satiety and energy expenditure.

The current understanding of the systems of the body and brain related to the control of energy intake, energy balance, and body weight has increased significantly since 2001. In the periphery, a variety of substances are believed to be involved in integrating physiological and behavioral systems directed towards energy homeostasis. The hormones insulin, leptin (secreted by adipocytes), and ghrelin (mainly secreted by the stomach) link the control of peripheral energy metabolism to the central nervous system (CNS) feeding behavior integrating unit by modulating short-term signals that determine meal initiation and termination as well as energy balance. The adrenal glucocorticoids (cortisol) predominantly influence the ingestion and metabolism of fat and carbohydrates. In addition to these hormones, the process of integrating metabolic information from the periphery in the CNS requires specialized functions of multiple brain areas. These include the lower brainstem, which integrates neural information between peripheral autonomic-endocrine organs and forebrain structures, the hypothalamus, which is closely linked to circulating metabolites and hormones, and the forebrain structures, such as the nucleus accumbens and amygdala, which integrate incoming information with various cognitive aspects of food.

The role of the hypothalamus in the process of relating hormones and metabolism to eating behavior has received considerable attention. A number of neuroendocrine systems have been identified in this structure and are believed to be involved in controlling appetite for the macronutrients, carbohydrate, fat, and protein. They also modulate metabolism and contribute ultimately to weight gain and obesity.

Stress-Eating

Many overweight people, particularly those with abdominal obesity, find that eating increases in situation of discomfort and also soothes their feelings of stress and anxiety. This phenomenon is usually called stress-eating or emotional-eating. While obese individuals may realize that in general they eat in response to stress/emotions, they are typically unable in the moment to stop themselves from eating. As will be discussed below, this type of stress-eating might be due to disruptions of central regulatory systems involved in the control of energy intake. In obese binge eaters, stress-eating occurs, and it is up to 100% specifically triggered by stress and negative moods (e.g., depression, anger, anxiety, etc.).

Stress-Eating and Obesity

It is not clear if stress-eating is secondary to being obese or vice versa, stress causing overeating, resulting in obesity. It is, however, clear that obese subjects have more body image distortion and distress about their weight than normal-weight subjects. Indeed, in our society being obese is so stigmatized that few individuals who are obese are able to feel good about their size. Negative body image and feelings of stress and anxiety about being obese may thus

serve as a trigger for stress-eating. In addition, stress and anxiety about being obese may work indirectly, increasing negative emotions such as depression and thus rendering the individual more susceptible to eating in response to stress. However, no studies have been reported indicating that obese subjects overeat by stress mechanisms leading to a maintained or escalated obese state.

Our knowledge of human obesity has progressed beyond the simple generalization that obesity is fully explained by inappropriate eating. Still, in more than 80% of the cross-sectional studies linking intake data to degree of obesity, dietary fat intake was associated with obesity. It is, however, important to recognize that human obesity is not a single entity, and subgroups are clearly more or less afflicted by stress. Realizing this distinction, it may be possible to examine how stress may be directly associated with an increased energy intake, and eventually with obesity.

Obesity with an increased number of adipocytes (i.e., hypercellular obesity) typically begins in early or middle childhood. In contrast, obesity with enlarged adipocytes (i.e., hypertrophic obesity) tends to correlate with truncal fat distribution, which in turn is an important predictor of the health hazards of obesity. Certainly, differences in clinical symptomatology favor a separation of obesity into two broad categories: visceral and nonvisceral obesity. These two forms seem to have different associations to stress. Central, visceral obesity is often associated with psychosocial and socioeconomic handicaps and excess use of alcohol and tobacco, as well as traits of depression and anxiety. Psychosocial stressors might induce eating as comfort food that may directly result in reduction of the negative effects of the chronic stressor in the nucleus accumbens. In addition, socioeconomic difficulties may make it hard to purchase nutritionally healthy food.

Mechanism of Stress-Eating

The expression of appetite is underpinned by a complex set of peripherally generated signals. Key hormonal appetite factors include insulin, cortisol, and leptin. Leptin is a protein that is synthesized exclusively in adipose tissue and moderates the drive to consume food. Within the neural circuitry numerous neuropeptides, both orexogenic (e.g., neuropeptide Y and melanocortin), and anorexigenic (e.g., proopiomelanocortin and corticotropin-releasing hormone) have been identified which stimulate and inhibit food intake.

Resistance to the appetite-suppressing effects of leptin is associated with common forms of obesity. Is it possible that stress factors associated with central, visceral obesity might be interfering with the leptin system? The human body holds many effector systems including the hypothalamic-pituitary-adrenal (HPA) axis for maintaining homeostasis. The secretory end product of the HPA axis, cortisol, is kept within an optimal range through the feedback action of cortisol interacting with neural control mechanisms. Distressing events or situations evokes prominent HPA system activation, and after long-term exposure, the HPA axis will eventually become dishabituated, resulting in a disruption of central regulatory systems. Together with insulin, the cortisol stimulate drive for and ingestion of comfort foods that may directly result in reduction of the negative effects of the chronic stressor in the nucleus accumbens, through stimulation of the anterior, more pleasure-associated part of this cell group, thus reducing the weight of the stress-stimulated posterior, more defensive part. Studies in rats have indicated that cortisol is directly involved in the sensitivity of the leptin regulation of satiety. When rats are adrenalectomized, their leptin sensitivity becomes high and feeding is inhibited. When corticosterone is replaced in a dose-dependent manner, food intake and body weight follow in parallel, and with excess corticosterone, obesity due to overeating appears. Furthermore, most animal models of obesity have enlarged adrenals and overproduction of glucocorticoids. Upon adrenalectomy, the somatic feature of obesity disappears. In normal-weight subjects, cortisol induces a dose-dependent increase in plasma leptin concentrations. In obese humans, food intake is clearly elevated after glucocorticoid administration in spite of increased levels of circulating leptin, indicating the presence of leptin resistance. Taken together it seems reasonable to assume that stress-eating is primarily caused by interference of stress-related cortisol secretion, with the normal signaling of satiety by the leptin system. The resulting deficient satiety signal causes overeating and finally the appearance of overweight and obesity.

Further Reading

Björntorp, P. and Rosmond, R. (2000). Obesity and cortisol. *Nutrition* **16**, 924–936.

Dallman, M. F., Pecoraro, N. C. and la Fleur, S. E. (2005). Chronic stress and comfort foods: self-medication and abdominal obesity. *Brain, behavior, and immunity* **19**, 275–280.

Gautier, J. F., Chen, K., Salbe, A. D., et al. (2000). Differential brain responses to satiation in obese and lean men. *Diabetes* **49**, 838–846.

Mastorakos, G. and Zapanti, E. (2004). The hypothalamic-pituitary-adrenal axis in the neuroendocrine regulation

of food intake and obesity: the role of corticotropin releasing hormone. *Nutritional Neuroscience* **7**, 271–280.
Rosmond, R., Dallman, M. F. and Björntorp, P. (1998). Stress-related cortisol secretion in men: relationships with abdominal obesity and endocrine, metabolic and hemodynamic abnormalities. *Journal of Clinical Endocrinology and Metabolism* **83**, 1853–1859.
Rosmond, R. (2005). Role of stress in the pathogenesis of the metabolic syndrome. *Psychoneuroendocrinology* **30**, 1–10.
Udden, J., Björntorp, P., Arner, P., et al. (2003). Effects of glucocorticoids on leptin levels and eating behaviour in women. *Journal of Internal Medicine* **253**, 225–231.
Zakrzewska, K. E., Cusin, I., Sainsbury, A., et al. (1997). Glucocorticoids as counterregulatory hormones of leptin: toward an understanding of leptin resistance. *Diabetes* **46**, 717–719.

Obsessive–Compulsive Disorder

R T Rubin
VA Greater Los Angeles Healthcare System, Los Angeles, CA, USA
B J Carroll
Pacific Behavioral Research Foundation, Carmel, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R T Rubin, volume 3, pp 77–82, © 2000, Elsevier Inc.

Introduction
Clinical Characteristics of OCD
Biological Characteristics of OCD
Brain Circuit Theories of OCD
Neuropsychological Testing in OCD
Neuroimaging Studies in OCD
Conclusions

Glossary

Antidepressants	Drugs that are used in the treatment of depression and other psychiatric disorders, including obsessive–compulsive disorder (OCD).
Basal ganglia	Groups of neurons deep in the brain that serve as integrative way stations for brain circuits.
Cerebral cortex	The outer layer of the brain, composed mainly of nerve cells (neurons).
Compulsions	Repetitive behaviors or mental acts that reduce anxiety or distress.
Neurotransmitters	Chemicals that carry messages from one nerve cell to another.
Obsessions	Intrusive and inappropriate recurrent and persistent thoughts.

Introduction

Psychiatric disorders are considered functional when there is no clearly understood central nervous system (CNS) pathophysiology underlying them. Thus, schizophrenia, the major affective illnesses, and anxiety disorders have been and are still considered functional psychiatric illnesses. This is an evolving nomenclature, however – as our knowledge about the pathophysiology of the brain in these disorders increases, the adjective functional conveys less and less meaning.

The brain substrates of the major psychiatric disorders are being examined in several ways. Complex genetic underpinnings are slowly being defined. Structural and functional imaging studies of the brain are providing some anatomical localization, limited by the resolution of the imaging techniques and the particular methodology of each study. Specific classes of drugs are known to be therapeutic for specific disorders, but the balance among chemical systems in the brain makes inferences about disturbance of a specific system in a particular psychiatric disorder difficult. Furthermore, no drug is curative of any psychiatric disorder. Obsessive–compulsive disorder (OCD) is one of the functional psychiatric illnesses that is giving way to an understanding of its pathophysiology in the brain. Stress can play a major role in the severity of OCD, and the illness itself can be very stressful to afflicted persons.

Clinical Characteristics of OCD

OCD is classified among the anxiety disorders. Its hallmarks are recurrent obsessions and/or compulsions severe enough to be time consuming (more than 1 h/day) or to interfere with occupational or social functioning. At some point, the individual must

recognize that these symptoms are excessive or unreasonable (this does not apply to children), and, if another major psychiatric disorder is present, the content of the obsessions or compulsions must be unrelated to it. Finally, the OCD must not be a direct result of a drug or an underlying medical condition.

Obsessions are recurrent and persistent thoughts, impulses, or mental images that are experienced as intrusive and inappropriate, that cause marked anxiety and distress, and that are not simply excessive worries about real-life problems. The individual recognizes that the mental intrusions are a product of his or her own mind (this distinguishes them from psychotic symptoms such as thought insertion) and attempts to neutralize them with other thoughts or actions (for example, compulsive rituals). The content of obsessions is most often violent (for example, the murder of one's child), sexual, contaminative (for example, touching unclean objects), or doubting (for example, worrying repeatedly about having performed a particular act).

Compulsions are repetitive behaviors or mental acts engaged in to reduce anxiety or distress. Common compulsive behaviors are hand washing, ordering of objects, and repetitive checking (for example, a locked door); common mental acts are repetitive counting, praying, and saying words silently. The frequency and severity of these obsessions and compulsions as necessary for a diagnosis are a minimum for these activities; not infrequently, many hours each day are spent in the throes of OCD, and compulsions may be physically damaging (for example, washing one's hands until the skin is raw). As mentioned, the individual knows these are alien thoughts and behaviors and is distressed by them, but is powerless to moderate them.

There appear to be different OCD symptom profiles across subjects. An analysis of scores on the Yale-Brown Obsessive-Compulsive Scale in more than 300 OCD patients showed that there were four relatively independent clusters of symptoms: obsessions and checking, symmetry and order, cleaning and washing, and hoarding. Other recognized clinical variants include obsessional slowness (for example, taking several hours to perform routine bathing, grooming, and dressing each day) and obsessional rumination, commonly with a hypochondriacal theme. The background personality of individuals with OCD tends to be rigid and perfectionistic. The end result of severe OCD can be impressively bizarre behavior: the housewife who must purchase new china whenever guests are invited; the adult son who cannot join his parents for the weekend because he spends 12 hours checking that his house has been properly locked; the hoarder whose house is stuffed with trash and newspapers to the point that it is declared a fire hazard. At present, there are no laboratory tests that are diagnostic of OCD.

The lifetime prevalence of OCD is 2–3%; thus, it is not rare. OCD occurs equally in men and women and usually begins in early adulthood, although childhood OCD is not uncommon. It may co-occur with other psychiatric illnesses such as major depression, other anxiety disorders, and eating disorders. There is a high incidence of OCD (30–50%) in patients with Gilles de la Tourette syndrome, which consists of severe tics (sudden, involuntary muscle movements) and involuntary utterances, occasionally vulgar in nature. Tourette syndrome is quite rare, but approximately 5% of OCD patients may have some form of it, and 20–30% of OCD patients have a history of tics. Indeed, one formulation considers that an obsession is a psychic tic. OCD affects both twins more often in monozygotic (single-egg) twins than in dizygotic twins, and there is a higher incidence of OCD in first-degree relatives (parents, siblings, and children) of both OCD and Tourette syndrome patients than in the general population. OCD patients in stressful life situations often have a worsening of their symptoms, for example, more frequent intrusive obsessional thoughts or more frequent compulsive behaviors such as repetitive hand washing.

Biological Characteristics of OCD

Neurotransmitters (for example, serotonin, norepinephrine, and dopamine) are chemicals in the brain that are released by neurons and carry a chemical message across a short space (synapse) to activate receptors on downstream neurons. An interesting treatment finding is that OCD responds best to drugs that are serotonin uptake inhibitors (SUIs) in the brain. These drugs (for example, clomipramine, fluoxetine, paroxetine, sertraline, and fluvoxamine), block the serotonin transporter, which recycles the neurotransmitter serotonin from the synapse back into the nerve cell that released it. By blocking transporter uptake, these drugs increase synaptic serotonin concentrations, at least in the short term. This implies that reduced serotonin neurotransmission may be a neurochemical basis for OCD, but clear evidence of this, by measurement of serotonin or its metabolites, transport mechanisms, or receptors, has not yet been forthcoming.

Complicating this perspective is recent evidence that chronic treatment with an SUI completely blocks the rise of extracellular brain serotonin that occurs in response to stress. Serotonin is mainly an

inhibitory neurotransmitter in the brain, especially in areas that are implicated in OCD, such as orbital prefrontal cortex and hippocampus. In these sites, serotonin reduces the firing of cortical pyramidal cells that are essential components of the cortico-striato-thalamo-cortical circuits now implicated in OCD and major depression. SUIs are most often used as antidepressants, and it is interesting that OCD and major depression, two illnesses with different symptom patterns (although occurring together in some patients), respond to treatment with the same drugs. The common factor is now considered to be that these disorders have a similar neuropathology that is expressed in functionally segregated neuronal circuits in the brain. Relief of OCD symptoms by SUIs is not complete (about 35% on average), and only about 50% of patients respond, suggesting that more than one brain neurotransmitter system may be involved. The addition of antipsychotic drugs (for example, haloperidol) that block dopamine receptors on neurons, thereby reducing dopamine neurotransmission, can be effective, certainly when there is a tic disorder associated with the OCD, and perhaps in cases refractory to SUI treatment. Other drugs useful in selected cases as adjuncts to treatment with SUIs are lithium (used most commonly to treat manic-depressive illness) and fenfluramine, a direct stimulator of serotonin receptors. Interestingly, dextroamphetamine, a dopamine and norepinephrine neurotransmission-potentiating drug, can reduce OCD symptoms in some patients.

Another aspect of OCD that supports an underlying brain abnormality is the presence of OCD-like disorders in other mammalian species, including subhuman primates, dogs, cats, horses, pigs, cows, sheep, bears, and parrots. Compulsive behaviors in these species include repetitive grooming to the point of skin damage; hair and feather pulling; tail chasing, pacing, whirling, bouncing, and somersaulting; sucking, bar biting, and lip flapping; compulsive sexual behaviors; aggression; checking; and territorial marking. Some equine and canine compulsive behaviors are familial and breed-specific, providing support for genetic factors in these conditions.

It cannot be determined which of these behaviors animals engage in to reduce anxiety or distress vs. which are distressing to the animal, but the animal is powerless to control them, the latter being a necessary criterion for the diagnosis of OCD in humans, as indicated previously. Some behaviors, such as pacing, whirling, and compulsive sexual activity, may be engaged in because a captive animal is deprived of necessary physical exercise and environmental stimulation, and these behaviors provide some degree of pleasure and relief from boredom. Nevertheless, some compulsive behaviors are pathological, such as acral lick in dogs, resulting in dermatitis, and avian compulsive feather picking. Several of these animal behaviors have been treated successfully with SUIs as well as with dopamine-blocking drugs.

The pharmacology of drugs useful in treating OCD points to which brain areas might be involved. Both the cerebral cortex of the frontal lobes (frontal cortex), which receives prominent serotonin connections from brain stem nuclei, and the basal ganglia, which receive dopamine connections from the substantia nigra, have been considered in OCD pathology. However, it should be noted that psychological treatments such as behavior therapy also can have a significant therapeutic effect, and brain imaging studies have shown that behavior therapy can induce changes in brain metabolism in OCD similar to those induced by drug therapy. Whether psychological therapy induces changes in neurotransmitter activity similar to those induced by pharmacotherapy is as yet unknown. In some instances of severe, incapacitating OCD that was unresponsive to all other treatments, stereotaxic cingulotomy (surgery that interrupts connections between the thalamus and the prefrontal cortex) has been effective. A significant minority of such patients experience gratifying clinical remission, but a partial response is equally common. Some 30–50% of patients, however, have little improvement from surgery.

Brain Circuit Theories of OCD

Brain circuit theories of OCD represent variations on the postulated involvement of certain areas of the frontal parts of the brain (frontal lobes) and their functional connections with deeper parts of the brain – the basal ganglia, thalamus, and other structures, including hippocampus and amygdala. There also are connections from these structures back to the cortex itself. In a general sense, the frontal lobes of the brain mediate executive functions, that is, an individual's decision-making and behavioral actions based on detailed evaluation of environmental demands, along with an appreciation of their historical context and a coordinated affective/emotional component. Particular areas of the frontal cortex mediate executive behavior, social behavior, and motivation.

The underlying theme of the brain circuit theories of OCD is that the lower, or orbital, frontal cortex is a generator of afferent impulses to the caudate nucleus, which is a part of the basal ganglia, and the caudate in turn serves as a filter or gating area. The caudate nucleus is considered to normally suppress extraneous thoughts, sensations, and actions, with little need

for control by the frontal cortex. Some investigators have hypothesized a primary defect in caudate function in OCD, which leads to an increased requirement for the frontal cortex to suppress and emotionally neutralize intrusive neuronal activity. Other investigators have viewed overactivity of the orbital frontal–thalamic circuit as the cause of compulsions in OCD, since compulsions represent unmodulated behavioral drive. The functional result is that the caudate nucleus cannot reasonably filter all the neuronal impulses sent to it from frontal cortical areas; some inappropriate or otherwise innocuous stimuli therefore are passed on to other structures, where they acquire excessive emotional meaning that results in inappropriate thoughts and actions.

Electroencephalographic (brain wave) studies of OCD subjects undergoing cognitive neuropsychological testing also support involvement of the frontal lobes, and functional neuroimaging studies suggest increased blood flow and metabolic overactivity of the orbital frontal cortex. Imaging studies of the caudate nucleus are less consistent in their findings. Thus, it is not possible at present to determine from imaging studies where in the frontal cortical/caudate nucleus/thalamus/frontal cortical circuit the primary defect resides. Structural lesions (for example, strokes and tumors) of both the frontal cortex and the basal ganglia have been associated with obsessions and stereotyped behaviors resembling compulsions. And, as mentioned, drug treatments of OCD do not particularly clarify the matter, since alteration of cortical serotonergic systems with SUIs often results in only partial relief of symptoms, and dopamine-enhancing drugs, which act on caudate nucleus dopamine systems, have some role in the management of OCD symptoms, especially when there is associated tic disorder.

Current theories emphasize hyperactivation of the lateral orbitofrontal loop in OCD, with overactivity of the orbital frontal cortex, the anterior cingulate cortex (ACC), and the caudate nucleus. As indicated previously, these brain structures also have been implicated in major depression and anxiety disorders. The ACC has been further implicated in OCD through correlative functional magnetic resonance imaging (MRI) studies of ACC activity and error checking. Compared to normal control subjects, patients with OCD have excessive error-related activation of the rostral ACC, which is correlated with the severity of OCD symptoms. One of the primary executive functions of the ACC is error detection or mismatch recognition; dysfunction in this brain region might well be associated with the inability of OCD patients to achieve cognitive closure. Failure to achieve cognitive closure is precisely what they demonstrate in their inability to terminate rituals and obsessions despite intellectual awareness that these behaviors are not justified. For the OCD patient, it is never okay: there is always the possibility that he or she has failed to rule out some contingency. This cognitive style is also apparent in formal testing of, for example, deductive and inductive reasoning: Patients with OCD have no problem with deductive reasoning, but they are significantly impaired on inductive reasoning tasks.

Neuropsychological Testing in OCD

Neuropsychological tests can tap some of the components of frontal lobe function, but it is difficult to implicate the involvement of a particular area of the frontal lobes in OCD from impaired performance on a specific neuropsychological test. Studies indicate deficits in OCD patients' ability to shift task set. This is consistent with theories of disrupted response feedback, in which repeated compulsive acts are required to determine their adequacy. An inherent need for high accuracy in a given individual may aggravate the tendency to lose response feedback, setting up the repetitive behavioral pattern.

Another issue is whether neuropsychological deficits or changes in brain metabolic activity appears first. Altered metabolic activity may result in the behavioral and neuropsychological test changes found in OCD, but it also may be true that the anxiety and need for absolute mastery caused by the illness itself underlie the increased cerebral metabolic activity noted on electroencephalogram (EEG) and functional neuroimaging studies.

Neuroimaging Studies in OCD

With the advent of neuroimaging techniques, it now is possible to probe the structure and function of the living human brain. Brain structure can be viewed with x-ray computed tomography (CT) and MRI. Nuclear medicine techniques such as single-photon and positron-emission computed tomography (SPECT and PET) allow the visualization and quantitation of regional cerebral blood flow, glucose metabolic rate, and neurotransmitter receptor occupancy, which are indirect probes of regional neuronal function. SPECT and PET use radiolabeled compounds as tracers; as these compounds decay, high-energy photons are emitted that are counted by external detectors. The distribution of the tracer molecules is computed, and cross-sectional images of the brain are created in which image brightness is proportional to the underlying physiological process being measured.

Studies of brain structure with CT and MRI in OCD patients have not demonstrated consistent

abnormalities. Most of these studies, however, focused on the caudate nucleus or lateral cerebral ventricles; less attention has been paid to the frontal lobes. Taken together, the findings suggest that structural abnormalities may be more associated with subtle neurological deficits accompanying OCD in some patients than with the disorder itself.

In contrast to the contradictory data from structural imaging studies, a consistent finding in SPECT and PET functional neuroimaging studies of OCD has been increased frontal cortical blood flow and glucose metabolism (hyperfrontality), although the precise cortical location has varied. The hyperfrontality does not seem to be a reflection of heightened anxiety in OCD patients. On the other hand, a relationship may exist between OCD hyperfrontality and serotonin neurotransmitter dysfunction. SUIs not only reduce OCD symptoms, but also alleviate the hyperfrontality, as shown by SPECT and PET.

A tendency toward lateralized right frontal hyperactivity in OCD also has been noted. Right frontal hypermetabolism in OCD is consistent with theories of frontal emotional lateralization. The right frontal cortex is considered to play a greater role in perception of and reaction to negative primary emotions and stimuli, whereas the left frontal cortex may be more active in response to positive stimuli and positive social emotions. The inability of OCD patients to suppress negative primary emotions and thoughts occurs along with their right frontal hyperactivity. A similar formulation of predominantly right prefrontal activity has been advanced for major depression. Suppression of right hemisphere cerebral activity, either with the Wada test (transient suppression of a cerebral hemisphere by barbiturate injection into the same-sided carotid artery) or as a result of lesions, is associated with decreased apprehension, euphoric mood, and minimization of negative emotions, while left hemisphere suppression produces feelings of depression, guilt, and worries about the future. With regard to basal ganglia function, neither structural nor functional neuroimaging studies strongly substantiate the hypothesis of a caudate abnormality in OCD. They do not rule out this possibility, but rather indicate that whatever abnormality might exist has not been consistently seen with current functional neuroimaging technology.

Conclusions

There is considerable variation in obsessive and compulsive symptoms, ranging from normal time-limited worry, through obsessive–compulsive personality disorder, to self-consuming OCD. The criteria for OCD represent a consensus statement among mental health professionals, based not only on the nature of the OCD symptoms but also on their frequency and interference with social and occupational functioning. OCD thus remains an arbitrarily defined syndrome, no matter how much expertise underlies the definition, so long as there are insufficient empirical data to validate its qualitative difference from obsessive–compulsive personality disorder, as well as from other major anxiety syndromes.

While animal models of OCD provide an interesting theoretical perspective, as noted earlier, this analogy also suffers from a certain nonspecificity. Repetitive behaviors such as pacing, whirling, hitting oneself, toe walking, and rocking, as well as increased anxiety, also are common symptoms of autism, a pervasive developmental disorder that begins in early childhood and persists into adult life. In autistic adults, dietary tryptophan depletion, which depletes brain serotonin, increases these repetitive behaviors, and treatment with SUIs decreases them. This suggests an involvement of brain serotonin neurotransmitter systems in autism similar to that in OCD.

It is interesting to speculate on the opposite of OCD, that is, excessive neglect in conducting one's activities of daily living. A large segment of society fails to adequately check their surroundings and to protect themselves from even known threats. Preventable accidents are a leading cause of mortality and morbidity, and epidemics of avoidable diseases such as AIDS are only too well known. This can be considered inappropriate denial of dangerous situations that have the potential to cause real harm to the individual. However, we approach inappropriate denial and risk taking only from a cognitive, educational standpoint, for example, trying to convince persons at risk to just say no. Perhaps our relative failure to make headway with such an approach suggests that we should take a fresh, psychobiological look at this aspect of behavior. For example, how is compulsive risk taking psychobiologically similar to or different from compulsive checking and hand washing? Psychologically they could be viewed as opposite ends of a spectrum, the former representing pathological self-neglect and the latter pathological self-protection. These and other issues remain the exciting areas of future research into OCD.

Acknowledgment

This work was supported by NIH grant MH28380.

See Also the Following Articles

Behavior Therapy; Depression and Manic-Depressive Illness; Selective Serotonin Reuptake Inhibitors (SSRIs).

Further Reading

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). (DSM-IV). Washington, D.C.: American Psychiatric Association.

Barr, L. C., Goodman, W. K., Price, L. H., McDougle, C. J. and Charney, D. S. (1992). The serotonin hypothesis of obsessive compulsive disorder: implications of pharmacologic challenge studies. *Journal of Clinical Psychiatry* **53**(Supplement), 17–28.

Chamberlain, S. R., Blackwell, A. D., Fineberg, N. A., Robbins, T. W. and Sahakian, B. J. (2005). The neuropsychology of obsessive compulsive disorder: the importance of failures in cognitive and behavioral inhibition as candidate endophenotypic markers. *Neuroscience and Biobehavioral Reviews* **29**, 399–419.

Cummings, J. L. (1995). Anatomic and behavioral aspects of frontal-subcortical circuits. *Annals of the New York Academy of Science* **769**, 1–13.

Denys, D., Zohar, J. and Westenberg, H. G. M. (2004). The role of dopamine in obsessive-compulsive disorder: preclinical and clinical evidence. *Journal of Clinical Psychiatry* (Supplement 14), 11–17.

Dodman, N. H., Moon-Fanelli, A., Mertens, P. A., Pflueger, S. and Stein, D. J. (1997). Veterinary models of OCD. In: Hollander, E. & Stein, D. J. (eds.) *Obsessive-compulsive disorder*, pp. 99–144. New York: Marcel Dekker.

Fitzgerald, K. D., Welsh, R. C., Gehring, W. J., et al. (2005). Error-related hyperactivity in the anterior cingulate cortex in obsessive-compulsive disorder. *Biological Psychiatry* **57**, 287–294.

Hoehn-Saric, R. and Benkelfat, C. (1994). Structural and functional brain imaging in obsessive-compulsive disorder. In: Hollander, E., Zohar, J., Marazziti, D. & Olivier, B. (eds.) *Current insights in obsessive-compulsive disorder*, pp. 183–211. New York: Wiley.

Jenike, M. A. (1995). Obsessive-compulsive disorder. In: Kaplan, H. I. & Sadock, B. J. (eds.) *Comprehensive textbook of psychiatry* (6th edn., pp. 1218–1227). Williams & Wilkins: Baltimore, MD.

Jenike, M. A. (2004). Obsessive-compulsive disorder. *New England Journal of Medicine* **350**, 259–265.

McGuire, P. K., Bench, C. J., Frith, C. D., et al. (1994). Functional anatomy of obsessive-compulsive phenomena. *British Journal of Psychiatry* **164**, 459–468.

Otto, M. W. (1992). Normal and abnormal information processing: a neuropsychological perspective on obsessive-compulsive disorder. *Psychiatry Clinics of North America* **15**, 825–848.

Pelissier, M. C. and O'Connor, K. P. (2002). Deductive and inductive reasoning in obsessive-compulsive disorder. *British Journal of Clinical Psychology* **41**(Part 1), 15–27.

Rubin, R. T. and Harris, G. J. (1999). Obsessive-compulsive disorder and the frontal lobes. In: Miller, B. L. & Cummings, J. L. (eds.) *The human frontal lobes: functions and disorders*, pp. 522–536. New York: Guilford.

Rubin, R. T., Villanueva-Meyer, J., Ananth, J., Trajmar, P. G. and Mena, I. (1992). Regional xenon 133 cerebral blood flow and cerebral technetium 99m HMPAO uptake in unmedicated patients with obsessive-compulsive disorder and matched control subjects. *Archives of General Psychiatry* **49**, 695–702.

Saxena, S., Brody, A. L., Schwartz, J. M. and Baxter, L. R. (1998). Neuroimaging and frontal-subcortical circuitry in obsessive-compulsive disorder. *British Journal of Psychiatry* **173**(Supplement 35), 26–37.

Zielinski, C. M., Taylor, M. A. and Juzwin, K. R. (1991). Neuropsychological deficits in obsessive-compulsive disorder. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology* **4**, 110–126.

Oklahoma City Bombing, Stress Effects of

C Piotrowski and S J Vodanovich
University of West Florida, Pensacola, FL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C Piotrowski, volume 3, pp 83–85, © 2000, Elsevier Inc.

Site of the Bombing
Effects on Survivors
Impact on Children
Emergency Personnel and Disaster Mental Health

Glossary

Access victims	Individuals who offered support and human assistance to survivors and families of victims, such as mental health professionals, nurses, and clergy.
Direct-contact victims	Emergency and rescue workers who searched the rubble for survivors and dead bodies, in addition to morgue personnel.
Direct-impact victims	Survivors who witnessed the bombing, death, and maiming of immediate victims firsthand.

Disaster mental health	A competency-based intervention model, based on disaster-related psychopathology, that addresses a community's mental health needs.
Intraself risk factors	A victim's idiosyncratic perception of the meaning and appraisal of a catastrophe and any premorbid factors that impact postdisaster reactions.
Mass-casualty event	The multidimensional stressors that may result in differential patterns of posttraumatic reactions, both short-term and long-term.
Self-efficacy adaptational coping	Survivors' struggles to maintain psychosomatic and social equilibrium.
Socioenvironmental risk factors	The degree of ecological and interpersonal destruction and intensity of exposure to morbidity.

Site of the Bombing

The bombing of the Alfred P. Murrah federal office building in Oklahoma City, Oklahoma, on April 19, 1995, remains the most devastating domestic terrorist act on U.S. soil. This tragic event impacted a close-knit community in America's heartland and tore at the fabric of our national sense of security. The bomb blast had a devastating impact; 168 lives were lost and nearly 1000 people were injured. Sadly, 19 children were killed and 57 were injured, some quite seriously. A majority (60%) of the fatalities were government employees. Because the event occurred in the morning, thousands of commuters were in the downtown area, in close proximity of the blast, heard the shattering explosion, and witnessed the calamity unfold. The rescue response was immediate, with hundreds of emergency workers, medical personnel, and law enforcement officials attending to the scene. During the first 48 h, the city's hospitals were overwhelmed with the critically wounded and exasperated family members while the bombing scene was amass with rescue teams in search of the survivors and victims. In a matter of minutes, Oklahoma City experienced a calamity of monumental proportions. Although the Oklahoma City bombing has many features in common with most disasters, the fact that the perpetrators were U.S. citizens added an additional stress factor to the readjustment of victims and the recovering community. The resolution of intense emotional concerns has had a central role in the healing process of not only the residents of Oklahoma City but also the nation as a whole. Perhaps this national consciousness is most evident in the poignant annual commemoration for all the victims of this tragedy.

Effects on Survivors

Disasters are usually classified as natural (e.g., hurricanes) or human-made (Three-Mile Island), but with the bombing in Oklahoma City, the American psyche was introduced to the ominous reality of disaster by intentional human strategy. In 1995, Parson coined the term Oklahoma City Bombing Stress Response (OKC-BSR), defined as (1) traumatic stress responses and (2) traumatically shattered meaning elements. In understanding victims' collective psychological responses and adaptation, Parson argued that ORC-BSR involves psychophysiological reactions to dissociation, anxiety symptoms, and acute stress disorder with a concomitant need to find meaning from a terrorist act. To address the overwhelming need for mental health services in the Oklahoma City area, a special program for individuals reporting psychological distress from the bombing (entitled Project Heartland) was established. Some 9000 residents received services such as crisis intervention and support groups. This program offered a model for investigating risk factors for posttraumatic stress disorder (PTSD) symptoms among the survivors of the trauma-related community stress.

Research studies on mental health symptomology and stress syndromes following the Oklahoma City bombing have been widespread in recent years. A study by Sprang, based on the premise that post-disaster reaction is a function of degree of exposure, found Oklahoma City residents reported high levels of PTSD and victimization symptomatology. Tucker and colleagues reported on 86 adults who sought assistance for stress related to the bombing 6 months after the tragedy. Reactions of nervousness to how others reacted at the time of the bombing accounted for approximately one-third of the variance in post-traumatic stress scores. These findings suggest that people who experience high anxiety during a traumatic event are subject to prolonged stress reactions over time. Pfefferbaum and colleagues reported in a series of studies on the long-term psychopathology and PTSD symptoms evident in both the survivors and the local population after the bombing. Survivors ($n = 37$) in the direct path of the explosion were assessed 17 months postdisaster. The results showed that 31% of the sample was designated as chronic PTSD. In another survivor sample ($n = 182$), 45% was diagnosed with postdisaster psychiatric disorders and 45% exhibited PTSD symptoms, particularly intrusive memories and hyperarousal, 6 months after the bombing. Furthermore, Pfefferbaum and colleagues, in a group of 84 individuals from the Oklahoma City community seeking support and mental health services, found a strong association between

PTSD and substance use behaviors such as smoking and alcohol intake. Such coping mechanisms have been found to exacerbate functional impairment in trauma/disaster victims. These studies on the long-term effects of the bombing confirm not only the immediacy of PTSD onset but also lingering stress response syndromes in populations with proximity to the disaster impact area. This calls for early intervention with long-term treatment options.

Because victims of a disaster may have had previous trauma, PTSD, or premorbid psychiatric conditions, they may be vulnerable to reacting pathologically to subsequent stressful events. In this regard, Moyers reported on the exacerbation of PTSD symptoms following the Oklahoma City bombing in veterans who were attending support groups.

Impact on Children

In recent years, the effects of major disasters on the mental health and emotional functioning of children has become a specialized area of research endeavor. Children are a particularly vulnerable group with the added stress of adapting to developmental milestones and dealing with emotional turmoil. In a sample of over 3000 middle school and high school students in the Oklahoma City school system, researchers found that children who knew someone who was killed by the blast (vicarious exposure) reported higher PTSD severity than did children who were not highly exposed to the traumatic event. Bereaved youths were more likely than nonbereaved peers to experience symptoms of anxiety and fear, changes in both school and home environments, and PTSD-type symptoms. Interestingly, girls reported significantly higher PTSD severity scores than boys. In addition, students who viewed extensive television coverage of the bombing were most negatively affected. These findings support prior research on the functional role that the defense mechanism of denial plays in the adaptation of children to the onerous impact of psychic trauma and disasters.

Emergency Personnel and Disaster Mental Health

A review of the literature on trauma provides conflicting evidence on the prevalence of stress reactions in emergency services personnel and rescue workers in the aftermath of disasters. Studies of firefighters and rescue personnel who responded to the bombing reported an unexpectedly low incidence rate (13%) for PTSD, even 3 years postdisaster. Body handlers were found to have negligible symptoms of depression or PTSD, and their overall mental health status improved over time. The resilience factors that aided in adaptation for these emergency personnel were identified as professional hardiness, job experience, and postdisaster debriefing.

Conversely, mental health workers in Oklahoma City reported chronic stress due to the increased workload for an extensive time period after the bombing. In fact, disaster mental health personnel were most prone to compassion fatigue and burnout. These findings illustrate the critical importance of professional education and psychosocial support initiatives. The implications for occupational stress, mental health, and organizational efficacy for emergency service workers cannot be ignored. Future research must focus on competency in conducting both disaster planning and postdisaster intervention.

See Also the Following Articles

Alarm Phase and General Adaptation Syndrome; Disaster Syndrome; Disasters and Mass Violence, Public, Effects of; Terrorism; Workplace Stress; Hurricane Katrina Disaster, Stress Effects of.

Further Reading

Allen, S. F., Dlugokinski, E. L., Cohen, L. A., et al. (1999). Assessing the impact of a traumatic community event on children and assisting with their healing. *Psychiatric Annals* **29**(2), 93–98.

Figley, C. R. (ed.) (2002). *Treating compassion fatigue*. New York: Brunner/Routledge.

Gordon, N. S., Farberow, N. L. and Maida, C. A. (1999). *Children and disasters*. Philadelphia: Brunner/Mazel.

Moyers, F. (1996). Oklahoma City bombing: exacerbation of symptoms in veterans with PTSD. *Archives of Psychiatric Nursing* **10**, 55–59.

North, C. S., Nixon, S. J., Shariat, S., et al. (1999). Psychiatric disorders among survivors of the Oklahoma City bombing. *Journal of the American Medical Association* **282**, 755–762.

North, C. S., Pfefferbaum, B., Tivis, L., et al. (2004). The course of posttraumatic stress disorder in a follow-up study of survivors of the Oklahoma City bombing. *Annals of Clinical Psychiatry* **16**, 209–215.

North, C. S., Tivis, L., McMillen, J. C., et al. (2002). Psychiatric disorders in rescue workers after the Oklahoma City bombing. *American Journal of Psychiatry* **159**, 857–859.

Parson, E. R. (1995). Mass traumatic terror in Oklahoma City and phases of adaptational coping. Part I: Possible effects on intentional injury/harm on victims' post-traumatic responses. *Journal of Contemporary Psychotherapy* **25**, 155–184.

Pfefferbaum, B. (2001). The impact of the Oklahoma City bombing on children in the community. *Military Medicine* **166** (supplement 2), 49–50.
Pfefferbaum, B. and Doughty, D. E. (2001). Increased alcohol use in a treatment sample of Oklahoma City bombing victims. *Psychiatry* **64**, 296–303.
Piotrowski, C., Armstrong, T. and Stopp, H. (1997). Stress factors in the aftermath of Hurricanes Erin and Opal: data from small business owners. *Psychological Reports* **80**, 1387–1391.
Piotrowski, C. and Dunham, F. (1983). Locus of control orientation and perception of "Hurricane" in fifth graders. *Journal of General Psychology* **109**, 119–127.
Reyes, G. and Elhai, J. D. (2004). Psychosocial interventions in the early phases of disasters. *Psychotherapy: Theory, Research, Practice, Training* **41**, 399–411.
Sprang, G. (1999). Post-disaster stress following the Oklahoma City bombing – an examination of three community groups. *Journal of Interpersonal Violence* **14**, 169–183.
Tucker, P., Dickson, W., Pfefferbaum, R., et al. (1997). Traumatic reactions as predictors of posttraumatic stress six months after the Oklahoma City bombing. *Psychiatric Services* **48**, 1191–1194.

Opioid System *See:* Opioids.

Opioids

J A Russell
University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J A Russell and A J Douglas, volume 3, pp 86–98, © 2000, Elsevier Inc.

Overview
Neuroendocrine Responses
Opioids and Stressor Processing
Opioids, Stress, and Reproduction
Conclusion

Glossary

β-Endorphin — A 31-amino-acid endogenous opioid peptide that acts mainly on δ- and μ-opioid receptors, synthesized from the proopiomelanocortin (POMC) gene. Cells expressing POMC are found in the arcuate nucleus of the hypothalamus, the corticotroph cells of the anterior pituitary, and the cells of the intermediate pituitary. Nerve fibers containing β-endorphin innervate the hypothalamus and project to other brain regions.

Agonist — Drug or exogenous or endogenous substance that binds to a receptor in a cell to induce a response.

Analgesia — The relief from pain; it may be local, spinal cord-mediated, or central.

Antagonist — A drug or endogenous substance that binds to a receptor in a cell-displacing bound agonist to prevent or reverse the effect of the endogenous ligand or agonist.

Dependence — The result of chronic exposure to an opiate drug that causes withdrawal signs to occur when the opiate exposure is interrupted; also called addiction.

Dynorphins — Endogenous opioid peptides that are 8 or 17 amino acids long and act mainly on κ-opioid receptors but also partly on δ- and μ-opioid receptors; synthesized from the preprodynorphin gene (also known as preproenkephalin B).

Endomorphins — Endogenous opioid peptides (discovered in 1997) that are 4 amino acids long and act selectively on the μ-opioid receptor; the gene or precursor is not known.

Enkephalins — Endogenous opioid peptides that are five amino acids long (Met^5 and Leu^5 enkephalins) and act mainly on δ- but also on κ- and μ-opioid receptors; synthesized from the preproenkephalin A gene. They are present in neurons widely distributed in brain, spinal cord, and adrenal medulla.

Euphoria — The conscious sensation of pleasure or feeling of well-being.

Limbic system Part of the brain comprising the amygdala, bed nucleus of the stria terminalis, hippocampus, septum, and nucleus accumbens that is concerned with memory, emotionality, and the processing of stressors.

Morphine An opiate alkaloid produced by the opium poppy that acts in the brain and peripherally to cause analgesia, euphoria, and reduced alarm; it is a relatively selective μ-opioid receptor agonist. Opium and heroin, a modified form of morphine, are widely abused for their pleasurable effects, but they induce tolerance and dependence (addiction).

Naloxone A general opioid receptor antagonist with greatest activity at μ-opioid receptors; it enters the brain from the circulation readily and reverses opioid agonist effects rapidly. It precipitates withdrawal symptoms in opiate addiction.

Nociceptin A peptide related to dynorphin that binds to ORL1 (OP4) receptors, which are related to κ-opioid receptors. It lacks the classic action of opioids, is not antagonized by naloxone, and has central pronociceptive properties; also called orphanin FQ.

Nociception Processes involved in perceiving pain following the stimulation of nociceptive nerve endings (nociceptors).

Opiate A drug derived from exogenous sources, originally the opium poppy, producing opium-like effects.

Opioid A drug or endogenously produced compound (e.g., opioid peptide) that acts through opioid receptors to produce effects that are antagonized by naloxone.

Opioid receptors Proteins in the cell membrane that confer sensitivity to opioids and opiates in the cells that express them; there are three, μ, δ, and κ. They have seven transmembrane domains, ligand-binding sequences projecting extracellularly, and a G-protein binding region intracellularly; a $G_{i/o}$-protein mediates the intracellular effects.

Orexigenic Being appetite stimulating; applied to an endogenous substance.

Pain threshold Often defined by the time a subject takes to respond to an applied painful stimulus of set intensity (e.g., in the tail flick test, which is applying radiant heat to the tail).

Plasticity Adaptation of an endogenous system so that its response to a particular stimulus or set of stimuli changes with repeated exposure or with a change in the physiological state of the individual.

Tolerance The effect of an opiate drug on receptor and postreceptor mechanisms in a cell so that the same dosage of a drug has a reduced effect and the dosage must be increased to maintain the same effects.

Overview

The soporific and analgesic properties of opium have been known to humans for thousands of years. Morphine, an alkaloid produced by the opium poppy, *Papaver orientale*, is the prototypic opioid. It and related synthetic opiates, such as heroin, have been widely used to control pain while reducing the state of alarm in those with severe injury or suffering the pain of acute myocardial infarction, advanced cancer, surgery, or childbirth. These drugs are widely abused because of their euphoric properties, but they induce tolerance and dependence. Sexual activity is reduced in opiate-dependent individuals. An overdose of morphine results in respiratory depression that can lead to death, but the potent antagonist naloxone acts rapidly to reverse the depression. Experimentally, these drugs have provided the tools to characterize opioid mechanisms. Pharmacological studies with various opiate drugs and antagonists identified three principal types of opiate receptor (μ, δ, and κ, according to their affinities for different agonists). These were found in the brain, spinal cord, posterior pituitary gland, gut, and other tissues. They are proteins, synthesized within the opioid-sensitive cells and inserted into the lipid plasma membrane so that the amino acid sequence that tightly binds the opioid ligand is exposed on the cell surface. The hypothesis that the expression of these high-affinity receptors for opiate drugs indicated the presence of endogenously produced ligands led to the discovery of the endogenous opioid peptides in the early 1970s. Three families were discovered: the enkephalins, β-endorphin, and the dynorphins. Subsequently, a peptide related to the dynorphins, nociceptin (or orphanin FQ), was discovered, but this acts on a different receptor (ORL1) that the other opioids do not act on. Naloxone is not effective at this receptor, and nociceptin lacks the classical actions of opioids. Indeed, in some paradigms it enhances pain perception. The three types of opioid receptors, although defined initially on the basis of their distinctive affinities for different opiate drugs and, in the early 1990s, by the characterization of their genes, in general do not selectively bind the different types of endogenous opioids. The enkephalins bind to μ- and δ-, β-endorphin binds to μ- and κ-, and the dynorphins bind to κ-receptors. However, opioid peptides (endomorphins 1 and 2)

selective for the μ-receptor have been identified in the brain. The enkephalins, β-endorphin, the dynorphins, and nociceptin are produced by the processing of higher-molecular-weight precursors, each translated from a distinct gene. However, each of the opioid precursor proteins contains opioid peptide sequences also present in the other two, indicating their common origin from an ancestral gene. Similarly, the opioid receptors are all members of the seven-transmembrane-domain G-protein-linked family, with considerable homology in their sequences. The G-protein binding part of the receptor is in the cytoplasmic loops. The intracellular G-protein links the receptor to the intracellular process that the opioid affects; this link may be direct (e.g., to an ion channel) or involve the inhibition of second-messenger (cAMP) generation. The G-protein is the $G_{i/o}$ type; it is inactivated by the pertussis toxin, and this is used as an experimental tool to determine whether a $G_{i/o}$-protein is involved in a particular opioid action.

Two principal types of opioid actions on neurons have been identified: presynaptic and postsynaptic. When opioids act presynaptically on nerve terminals, they inhibit the entry of Ca^{2+}; because such Ca^{2+} entry is essential for the release of a neurotransmitter into the synaptic cleft, opioids inhibit release, reducing the effectiveness of the synapse. If the opioid-inhibited presynaptic neuron produces an excitatory transmitter, then the postsynaptic neuron is inhibited; conversely the presynaptic opioid inhibition of an inhibitory neuron excites the postsynaptic neuron. Opioids can also act postsynaptically on the dendrites or cell bodies of neurons, inhibiting the processes that generate action potentials and, thus, indirectly inhibiting the release of the neurotransmitter from the terminals of these neurons. The mechanism of this type of opioid inhibition involves the activation of K^+ channels in the plasma membrane; the outflow of K^+ then hyperpolarizes the neuron, making it less excitable. This mechanism also contributes to opioid actions on nerve terminals. Nociceptin has similar modes of inhibiting neural activity. The actions of opioids reverse when the opioid diffuses away or is broken down by local proteases into peptides that have little affinity for the opioid receptors. The combination of an opioid receptor with an endogenous peptide leads to the internalization and activation of other intracellular regulatory mechanisms that blunt further opioid action. It seems that these mechanisms do not follow the occupancy of the μ-receptor with morphine, and this may underlie the addictive properties of morphine.

Plasticity is a feature of opioid systems. This is evident in the increased or decreased expression of opioid receptors or opioid peptides, which may be a result of the altered activity of their genes. The expression of the preproenkephalin and preprodynorphin genes is regulated by promoter sequences responding to intracellular messengers. The pro-opiomelanocortin (POMC) gene (which encodes β-endorphin) has a glucocorticoid receptor-response element that mediates the inhibition of expression in the anterior pituitary but mediates its stimulation in β-endorphin neurons. Opioid receptor expression can change because of redistribution within a neuron or altered turnover at the plasma membrane.

The principal actions of opiate drugs on mental state and pain perception suggest that the activation of endogenous opioid mechanisms during stress moderate the processing of stressors and the responses to them. There are several lines of evidence for this. First, the opioid peptides and their receptors are found, at appropriate sites, to influence reactions to stress. Second, the actions of opioid antagonists reveal activity in endogenous mechanisms; the effects of exogenous opioids, and the consequences of absence of opioid peptides or their receptors in mutants, all point to important roles for endogenous opioids in stress. Third, the loss of responses suspected to be due to endogenous opioid action after chronic morphine treatment, which produces tolerance to opioid actions at μ-receptors, indicates a cross-tolerance to an endogenous opioid action. It is clear that endogenous opioid mechanisms are involved in modulating reactions to stress at many levels. For physically painful stimuli, endogenous opioids modulate processing in the nociceptive pathways. Opioids are present in the brain-stem nuclei concerned with stressor processing and, similarly, in the limbic system. Opioids act in parts of the brain concerned with pleasure and reward, and here they affect appetite. Opioid peptides are coexpressed in the neuroendocrine neurons that mediate hormonal stress responses, and these neurons can be targets for opioid peptides. In the periphery, endogenous opioids modify the activity of immune cells and modulate inflammatory pain by actions on nociceptive nerve terminals. Opioids impinge on the neuroendocrine neurons regulating gonadal activity.

The major sources of endogenous opioids are the anterior pituitary gland and the medullae of the adrenal glands, which can both secrete opioid peptides (β-endorphin and enkephalins, respectively) into the circulation when stimulated; the brain and spinal cord, where opioids are produced by specific neurons; and the cells of the immune system, when activated. Opioids produced by neurons (β-endorphin, enkephalins, or dynorphins) act locally in the immediate vicinity of the site of release from the terminals of

the processes of the neurons. Enkephalins and dynorphins are commonly produced in neurons also producing other transmitters, whereas β-endorphin neurons also produce other products of the POMC gene. Opioids from neurons generally inhibit the activity of other neurons or act back on the neuron terminals from which they have been released, inhibiting further the release of transmitters. The opioids (principally β-endorphin) produced by immune cells also act locally at the site of inflammation. For opioids secreted into the circulation, there is a problem about their site(s) of action. The concentrations of opioid peptides in the circulation may not reach levels sufficient to activate the receptors in distant tissues, and blood-borne opioids may not penetrate the tissues to reach cells bearing opioid receptors. In particular, the tight junctions between the lining cells of the blood capillaries in most regions of the brain (the blood–brain barrier) prevent the free diffusion of peptides into the brain. There are transport mechanisms in the choroid plexus for selectively transporting certain peptide messengers (e.g., leptin) from blood into the cerebroventricular system. However, there is little transfer of β-endorphin from the blood into cerebrospinal fluid, and it does not then permeate the brain tissue. Enkephalins produced in the brain are removed partly by being transported into the blood, which seems to preclude their being transported in the opposite direction. Dynorphins are not transported into the brain. Furthermore, in experimental animals, after the removal of the pituitary and adrenal glands, thus eliminating these sources of opioid peptides, pain control during stress can persist; this shows the importance of opioids produced within the brain or spinal cord. Thus, it is unlikely that the activity of neurons in the brain can be directly affected by opioid peptides circulating in the blood. This is not true of the opiate drugs, such as morphine, which enter the brain and have major actions on receptors and on neurons inside the blood–brain barrier, or of antagonists, particularly naloxone, which enters the brain readily and reverses the effects of an opiate drug overdose very rapidly. A modified form of naloxone (naloxone methylbromide), with an altered charge, does not cross the blood–brain barrier; this is used experimentally to distinguish the activity of opioids within the brain from the activity of those outside.

In general, the activation of endogenous opioid mechanisms by stress reduces the perception of pain; modulates autonomic (sympathetic nervous system) and neuroendocrine mechanisms that bring about metabolic adjustments to meet the stress; stimulates eating, especially of highly palatable foods during stress; and inhibits energetically costly activities that are not essential for immediate survival, especially reproduction. This article reviews the mechanisms.

Neuroendocrine Responses

Opioids and the Hypothalamic-Pituitary-Adrenal Axis

Endogenous opioid and opioid receptor expression in the hypothalamic-pituitary-adrenal system The defining neuroendocrine outcome of stress is the release of cortiocotropin releasing hormone (CRH) and vasopressin from the terminals of parvocellular paraventricular nucleus (PVN) neurons in the median eminence, causing increased secretion from the anterior pituitary gland of adrenocorticotropic hormone (ACTH), which stimulates the secretion of glucocorticoids from the adrenal cortex. Endogenous opioids are cosynthesized and/or have actions at all these locations to modulate responses to stress (**Figure 1, i, ii**). In the paraventricular nucleus, the parvocellular CRH neurons also synthesize enkephalins from the preproenkephalin A gene and express opioid (at least, κ-) receptors. Preproenkephalin A gene expression is rapidly stimulated in these neurons in response to stress. Endomorphins 1 and 2 have also been described in neurons and fibers in the paraventricular nucleus. There are opioid receptors in the median eminence that mediate local regulation at the level of the terminals secreting releasing factors into the hypothalamo-hypophysial portal system. Opioids acting here could be enkephalins coreleased with CRH or secreted by the adrenal medulla or by β-endorphin from the anterior pituitary because the blood–brain barrier is deficient in this region. Corticotrophs in the anterior pituitary synthesize both ACTH and β-endorphin from the POMC gene, and they are cosecreted; however, much of the β-endorphin may be acetylated and, consequently, inactive. This changes in pregnancy (the placenta produces POMC in women). ACTH release from the corticotrophs is not directly influenced by any opioid released locally or carried in the portal blood from neuroendocrine neurons. The adrenal cortex does not produce opioid peptides in the rat, guinea pig, human, or bovine, but in the rat it does show specific receptors for β-endorphin and dynorphin in the zona fasciculata-reticularis. However, the adrenal medulla in many species (rat, mouse, human, guinea pig, horse, rabbit, and bovine) expresses preproenkephalin A and POMC-derived opioids. These appear to influence the adjacent cortex locally; enkephalins and dynorphins are generally stimulatory in the rat and endorphins are generally inhibitory to glucocorticoid secretion in both the rat and human. Adrenal

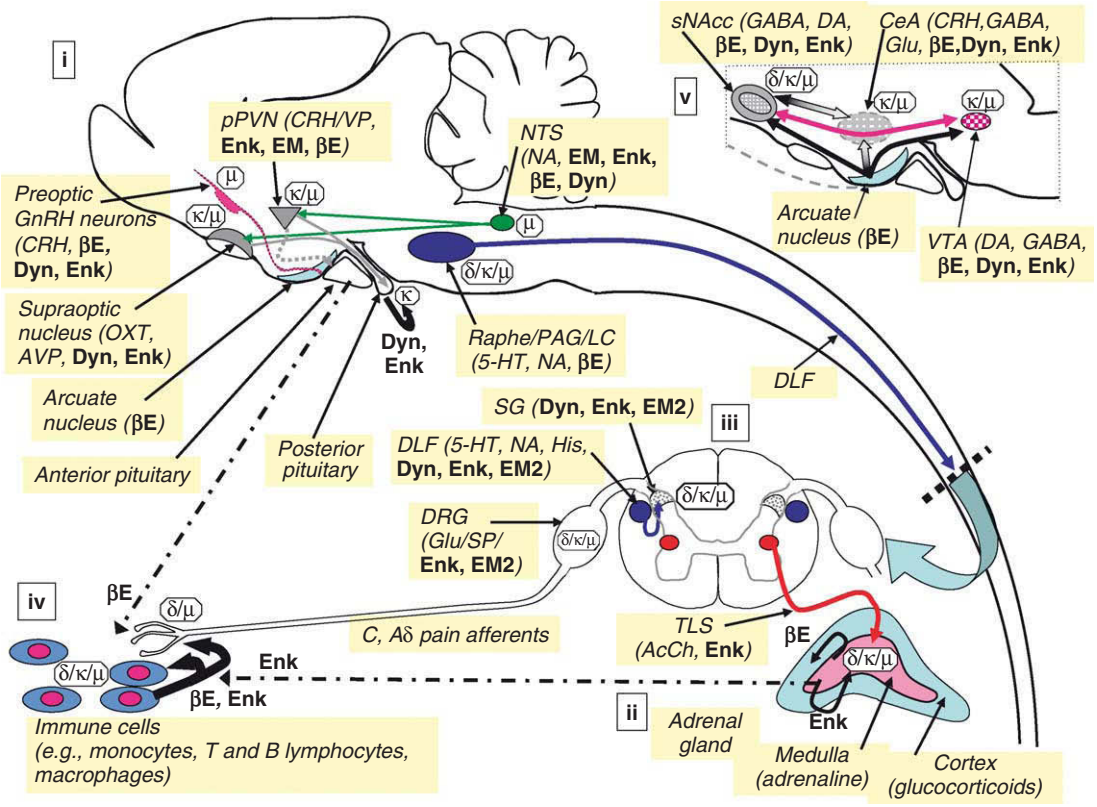


Figure 1 The distribution of opioid peptides and opioid receptors in key elements of central and peripheral structures that organize stress responses. The elements illustrated (not to scale) are (i) neuroendocrine brain and pituitary gland (sagittal section), (ii) adrenal gland, (iii) spinal cord pain pathways (transverse section), (iv) immune cells, and (v) mesolimbic reward circuit. For each element, the predominant types of opioid receptor and the type of opioid peptide (in parentheses in boldface) present or produced are indicated. Note that β-endorphin-producing neurons in the brain are located in the arcuate nucleus (with a few in the nucleus tractus solitarius) and project to many brain regions (as in (v)). Endomorphin-producing neurons are localized in the hypothalamus and nucleus tractus solitarius. Enkephalin- and dynorphin-producing neurons are more widely distributed. Other transmitters in the key central nervous system areas are indicated (in parentheses in italics). Neural pathways modulated by opioids are indicated by straight or long curved arrows; short curved arrows indicate local opioid action; dot-dash lines indicate opioids carried in blood. 5-HT, serotonin; βE, β-endorphin; δ, κ, and μ, δ-,κ-, and μ-opioid receptors; AcCh, acetylcholine; C, Aδ, C- and Aδ-fiber; CeA, central nucleus of amygdala; CRH, corticotropin releasing hormone; DA, dopamine; DLF, dorsolateral funiculus; DRG, dorsal root ganglion; Dyn, dynorphin; EM, endomorphin; Enk, leu-, met-, or extended enkephalin; GABA, γ-aminobutyric acid; Glu, glutamate; GnRH, gonadotropin releasing hormone; His, histamine; LC, locus coeruleus; NE, norepinephrine; NTS, nucleus tractus solitarius; OXT, oxytocin; PAG, periaqueductal gray; pPVN, parvocellular paraventricular nucleus; Raphe, raphe nucleus; sNAcc, shell of nucleus accumbens; SP, substance P; TLS, thoracolumbar sympathetic outflow; VP, vasopressin; VTA, ventral tegmentum.

corticosteroid secretion is increased after naloxone in the human independently of a change in ACTH secretion.

Endogenous opioid and opioid receptor expression in inputs to the hypothalamic-pituitary-adrenal axis Neural inputs to the hypothalamic-pituitary-adrenal (HPA) axis that contain opioids include β-endorphin nerve terminals of arcuate nucleus neurons (adjacent to the median eminence), which traverse the median eminence and innervate the hypothalamus, including the paraventricular nucleus. CRH neurons in the paraventricular nucleus are contacted by β-endorphin terminals. Thus, β-endorphin can act on both the cell bodies of CRH neurons and their terminals in the median eminence to influence CRH and vasopressin and/or opioid release at these sites. Endomorphin 2-expressing cell bodies have been located in the hypothalamus, near the third ventricle. In addition, opioids (enkephalins and dynorphins) are synthesized in the neurons of the nucleus of the tractus solitarius (NTS) and ventral medulla in the brain stem, colocalized in noradrenergic neurons, among others; these project to the hypothalamus, including to the paraventricular nucleus and median eminence (Figure 1, i). NTS neurons coexpressing prodynorphin or proenkephalin A mRNA are activated during seizures and other stressors, increasing their expression of prodynorphin mRNA after seizures. The limbic brain regions (the

amygdala; septum; bed nucleus of stria terminalis, BNST; and hippocampus) express opioid receptors and receive opioidergic inputs. Locus coeruleus noradrenergic neurons, important in arousal and attention in response to stressors, receive stimulatory CRH and inhibitory enkephalin-containing terminals, giving the basis for the activation and then the suppression of the activity of these noradrenergic neurons with an acute stressor. After chronic morphine exposure, the effect of CRH is enhanced, and this may exemplify how opiate addiction sensitizes to stressors and how stress predisposes to opiate craving. The amygdala has μ-opioid receptors in both rat and human, and it expresses endomorphin 2 in the rat; μ-, δ-, and κ-opioids all stimulate neurons in the amygdala in the guinea pig. Naloxone prevents stress-induced changes in long-term potentiation in the hippocampus. The septum expresses μ-opioid receptors. Enkephalin-containing neurons in the BNST are contacted by noradrenergic axon terminals and are activated by stressors. Nociceptin and its receptor (ORL1) have been described in both the amygdala and the hippocampus. δ-Receptors are not found in the paraventricular nucleus, pituitary, or adrenal cortex but may weakly contribute to regulating basal HPA secretion, perhaps via distant inputs. Thus, both brain-stem and limbic system neurons involved in mediating the activation of paraventricular nucleus CRH neurons by stressors either produce opioids or can be affected by them.

Endogenous opioids and hypothalamic-pituitary-adrenal stress responses The major effects of opioids on HPA axis stress responses are in the brain. Acutely administered opioids generally act in the hypothalamus to inhibit CRH secretion and hence ACTH secretion in most species. In the sheep, pig, and human, endogenous opioids inhibit basal and stress-stimulated HPA axis activity; and in the horse, opioids are inhibitory but naloxone increases ACTH secretion without altering CRH or vasopressin release. However, in the rat, opioids are stimulatory to the HPA axis. Thus, the release of CRH is increased following the central application of an opioid peptide (e.g. β-endorphin) *in vivo* or *in vitro,* and the effect is blocked by naloxone. The activation of endogenous opioid peptide mechanisms impinging on CRH neurons during stress (including emotional and physical stressors) is indicated by a reduced ACTH response after treatment with the antagonist naloxone. However, there are conflicting reports of tonic opioid-mediated inhibition. The subtle change of a single nucleotide polymorphism (SNP) in the rhesus monkey μ-opioid receptor gene, which increases the affinity of the receptor for β-endorphin, is associated with reduced basal cortisol secretion. Surprisingly, whereas the central administration of morphine, a μ-receptor agonist, increases corticosterone secretion in the rat, a selective endogenous μ-receptor agonist (endomorphin-1 or -2) does not. In humans, naloxone increases basal ACTH secretion in most subjects and has a greater effect in patients with depression than in patients with Cushing's disease. A naloxone test has been suggested as an aid to distinguishing these conditions, on the basis that naloxone acting in the brain will further enhance CRH release, and hence ACTH secretion, in depressed patients with hypercortisolism but not in patients with Cushing's disease. The different opioid effects on the HPA axis may be the result of different opioids acting through different types of opioid receptor, such that μ-opioids activate and κ-opioids inhibit, and at different sites, depending on the stimulus conditions. The balance of these different actions may even change the evident direction of action from stimulation to inhibition (or vice versa) and account for species differences. Such a change has been described in late pregnancy in the rat, when endogenous opioid inhibits the activation of CRH neurons in response to several stressors by inhibiting the release of norepinephrine in the paraventricular nuclei.

The opioid enhancement of HPA axis secretory responses in rats shows plasticity during chronic stress, aging, and pregnancy/parturition. Because chronic stress increases the expression of the preproenkephalin A gene in parvocellular PVN neurons, their release of enkephalins is expected to increase. Furthermore, the neurons in the parvocellular PVN that produce enkephalin and CRH continue to be activated by an acute stressor after chronic stress. This suggests an important role for the coproduced enkephalin in mediating or modulating the HPA stress response after chronic stress. The opioid regulation of stress responses seems to decline with age because naloxone has a reduced effect, both on HPA axis responses and on stress-induced analgesia. During parturition in the rat, the usual opioid enhancement of secretory responses switches to inhibition, restraining HPA axis secretion. This follows an attenuation of HPA axis stress responses in late pregnancy, partly due to a lack of endogenous opioid enhancement. However, arcuate nucleus POMC expression and hypothalamic β-endorphin content increase in pregnancy in the rat (β-endorphin concentration in the blood also increases in both rat and human during pregnancy). Furthermore, β-endorphin can reduce CRH expression in the PVN and reduce POMC expression and ACTH secretion in the pituitary after stress. Thus, although it is possible to define the importance and characteristics of opioid actions at multiple sites in

the neural networks regulating the HPA axis, the net effect on the output is not easy to predict. The experimental finding that chronic prenatal exposure to morphine can dampen the ACTH response of the offspring to stress when they become adult indicates that the baby born to an opiate-addicted mother may be programmed to respond less to stress.

In summary, endogenous opioid mechanisms are activated by stress and modify HPA axis responses to stress at several levels in the axis. The activity of these opioid mechanisms is not fixed, and this contributes to altered HPA responses to stress, for example, as a result of reproductive activity or chronic disease. β-Endorphin is secreted into the circulation when ACTH is secreted and may have distant actions through opioid receptors outside the brain.

Opioids in the Hypothalamic-Neurohypophysial System

Oxytocin is secreted into the blood from the nerve terminals in the posterior pituitary gland of magnocellular neurons that have their cell bodies in the supraoptic nucleus and PVN in the hypothalamus. These neurons, and hence oxytocin secretion, are stimulated by emotional and physical stressors in the rat (but generally not in other species), in which vasopressin is the posterior pituitary stress hormone. Although oxytocin may act directly on the adrenal cortex to enhance basal corticosteroid secretion, its role in the periphery in coping with stress remains unclear. Oxytocin secretion is inhibited by opioids. In the rat, vasopressin is secreted into the blood (from magnocellular neurons adjacent to oxytocin cells) in response to noxious stimuli and probably mediates cardiovascular compensation as well as conserving water by its renal actions; endogenous opioids can restrain the vasopressin response. Other stressors, such as conditioned fear, inhibit vasopressin secretion, but opioids are not responsible for this.

Opioid peptides and opioid receptor expression in the hypothalamic-neurohypophysial system Oxytocin neurons coexpress the endogenous opioids, enkephalins and dynorphins, in both the supraoptic nucleus and PVN. These opioids are costored in the nerve terminals in the posterior pituitary and are coreleased with oxytocin on stimulation, and thus they can autoregulate oxytocin secretion into the blood (**Figure 1, i**). Opioids originating from the nearby anterior and intermediate lobes of the pituitary are evidently not important. Endogenous opioids also come from the adjacent vasopressin neurons, which coexpress and secrete dynorphins from their nerve terminals in the posterior pituitary and from inputs to the oxytocin neuron cell bodies from the arcuate nucleus (β-endorphin) or brain stem (enkephalins). The magnocellular cell bodies of oxytocin neurons in the supraoptic nucleus and PVN have κ- and μ-opioid receptors, but their nerve terminals in the posterior pituitary have only κ-opioid receptors; vasopressin neurons are sensitive to κ-opioids only. Although Met^5- and Leu^5-enkephalins are not active at κ-receptors, the extended forms Met-enkephalin-Arg^6Phe^7 and Met-enkephalin-$Arg^6Gly^7Leu^8$ are; these are produced by oxytocin neurons, and like dynorphins they can inhibit stimulated oxytocin secretion from the posterior pituitary. δ-Opioid receptors are not evident in the hypothalamic-neurohypophysial system (HNS) and δ-acting opioids are not important in its regulation. Inputs to the magnocellular supraoptic nucleus and PVN arise from other areas of the hypothalamus, the brain stem, the circumventricular organs, the BNST, and the olfactory bulb. Although most of these regions express opioid peptides and opioid receptor binding, and opioids could affect the activity of these inputs, their role in oxytocin and vasopressin responses to stress is not known.

Endogenous opioids in hypothalamic-neurohypophysial system responses to stress Oxytocin neuron activity and secretion are inhibited by exogenous or endogenous opioids under stimulated conditions in males and females. Thus naloxone increases and κ- or μ-opioids inhibit oxytocin secretion in response to parturition, suckling, hyperosmotic stimuli, and stress (e.g., forced swim, restraint, surgery, or white noise), the major physiological conditions that increase oxytocin secretion. In the supraoptic nucleus, opioids can act directly on the oxytocin neurons, inhibiting their discharge of action potentials via κ- or μ-opioid receptors, and act presynaptically on μ-opioid receptors to inhibit the release of neurotransmitters from brain-stem inputs. However, the inhibition of oxytocin secretion by endogenous opioids is so far known to occur only at the oxytocin nerve terminals in the posterior pituitary, except in pregnancy. The mechanisms of opioid action here are similar to those on other nerve terminals. Opioid antagonists such as naloxone increase oxytocin secretion into the blood in response to stressors in the rat. Researchers believe that κ- and μ-opioid receptors are differentially involved in regulating responses to different stressors in the rat; thus, endogenous μ-opioids regulate oxytocin secretory responses to immobilization, whereas endogenous κ-opioids regulate responses to intraperitoneal hypertonic saline. The μ-opioid receptor actions are likely to be in the brain rather than the posterior pituitary, but the

sites of action of endogenous opioids during stress are unknown.

There is plasticity in the opioid regulation of the HNS. Acute hyperosmotic stress induces the secretion of oxytocin and vasopressin and upregulates proenkephalin (weakly) and prodynorphin gene expression in the supraoptic nucleus. However, chronic stress in rats more strongly upregulates preproenkephalin A gene expression in supraoptic neurons. In pregnancy, oxytocin secretory responses to stress are attenuated more strongly by endogenous opioids. A central site of opioid inhibition is likely at this time because there is a downregulation of κ-opioid receptors and a decreased sensitivity to κ-opioids in the posterior pituitary. The upregulated endogenous opioid mechanism in pregnancy involves increased proenkephalin A and μ-opioid receptor mRNA expression in the brain-stem input and, possibly, increased β-endorphin action from arcuate nucleus neurons. The upregulation of endogenous opioid inhibition of oxytocin neurons in pregnancy is driven by neuroactive steroid products of progesterone metabolism in the brain. The more powerful opioid restraint of oxytocin neurons that emerges in pregnancy acts to restrain the responses of oxytocin neurons to a range of stressors, including interleukin-1β. This opioid-induced resistance of the oxytocin system to stressors, and especially immune signals, in pregnancy may be important in preventing premature activation of parturition. Paradoxically, stress slows parturition by decreasing oxytocin secretion in a naloxone-reversible way. Stress may still have a stimulatory effect on oxytocin secretion in parturition, but may disrupt the pulsatile pattern of secretion that is essential for parturition (see **Oxytocin**).

In the rat, vasopressin secretion from the posterior pituitary into the blood is restricted to selected stressors – for example, to electrically applied foot shock, surgical stress, and noxious stimuli but not to restraint, forced swimming, noise, and novel environment exposure. Opioids generally inhibit magnocellular vasopressin secretion via central κ-opioid receptor mechanisms in rats and humans. Partly, this reflects the autoregulatory action of dynorphin, which is evidently released by the dendrites of vasopressin neurons during periods of electrical activity (phasic bursts) and, acting through κ-opioid receptors on the neurons, eventually terminates each burst. Otherwise, the effects of endogenous opioids on stress-induced vasopressin release are dependent on the stressor. For instance, naloxone augments the vasopressin secretory response to foot shock. Endogenous opioids are not responsible for the lack of vasopressin secretion in response to stressors such as forced swimming or restraint in the rat and this does not change in altered physiological conditions such as pregnancy.

In summary, for posterior pituitary hormone secretion, and on stimulated oxytocin secretion in particular, endogenous opioids have powerful moderating actions at several levels under stressful conditions. These actions can be seen as contributing to a successful outcome of pregnancy.

Opioids and the Adrenal Medulla

The stimulation of the sympathetic nervous outflow from the spinal cord drives the secretion of epinephrine from the adrenal medulla chromaffin cells as part of the flight-or-fight response to stress. The adrenal medulla is an important site of opioid modulation of stress responses.

Endogenous opioid and opioid receptor expression in the adrenal medulla Soon after the discovery of the endogenous opioids, it was recognized that the adrenal medulla produces abundant opioids. In particular, proenkephalin A is expressed by chromaffin cells, which produce the catecholamines, epinephrine, and norepinephrine. These are secreted into blood when chromaffin cells are excited through nicotinic receptors by their cholinergic preganglionic sympathetic innervation from the spinal cord, that is, in acute stress (**Figure 1, ii**). Thus, the phenylethanolamine *N*-methyltransferase (the enzyme synthesizing epinephrine) and preproenkephalin A genes are expressed in the same cells. The preproenkephalin A protein is processed by a serine protease to yield Met^5- and Leu^5-enkephalins but also the extended peptides Met-enkephalin-Arg^6Phe^7 and Met-enkephalin-Arg^6-Gly^7Leu^8. Enkephalins and epinephrine are packaged in the same secretory vesicles, and the stimulation of the chromaffin cells with acetylcholine or nicotine increases the secretion of both epinephrine and enkephalins into the blood, with the possibility of distant opioid actions. Such excitation increases the expression of the preproenkephalin A gene and the subsequent synthesis of enkephalins, so the secretion of opioid can be sustained during continual stimulation. Pro-inflammatory cytokines stimulate or enhance the release of enkephalins from the chromaffin cells, indicating that production is increased with infection. The capacity of adrenal chromaffin cells to produce opioids is demonstrated by the analgesia caused by fragments of adrenal medulla implanted adjacent to the dorsal horns of the spinal cord or in the periaqueductal gray (PAG) in experimental models of chronic pain. POMC is also expressed. Only a few adrenal medullary cells express the prodynorphin gene, except when these cells become tumorous (pheochromocytoma), when dynorphin is

abundantly produced. Nociceptin is also produced. There are opioid receptors in the adrenal medulla, κ-opioid receptors are present on the nerve fiber tracts and in the epinephrine-producing chromaffin cells, δ-opioid receptors are present on the norepinephrine-producing cells, and μ-opioid receptors are present at a low level. The adrenal medulla thus provides a microcosm to investigate the pre- and post-synaptic actions of opioids on neurons.

Endogenous opioids and adrenal medulla stress responses A presynaptic action of opioids has been shown with *in vivo* microdialysis of the adrenal medulla, revealing the inhibition by μ- and δ-opioid agonists of stimulated acetylcholine release. The cholinergic nerve terminals probably release enkephalin when stimulated, causing autoinhibition at these terminals as well as inhibition by enkephalin released from the chromaffin cells (**Figure 1, ii**). Exogenous opioids inhibit catecholamine (norepinephrine) release from chromaffin cells induced by splanchnic nerve stimulation or by nicotine, but this is not reversed by naloxone, and indeed opioid antagonists have a similar effect, so classic opioid receptors do not seem to mediate this action on the strong stimulation of the chromaffin cells. These actions of both opioid agonists and antagonists involve the inhibition of nicotine-stimulated Ca^{2+} entry, essential for exocytosis of the contents of the secretory granules. However, other studies at the level of individual cells show clear stereospecific, opioid receptor-mediated actions. Opioids released by chromaffin cells act through the δ- or μ-opioid receptors on these cells to inhibit secretion by reducing the opening of voltage-activated non-L type (N or P/Q) Ca^{2+} channels through a pertussis-toxin-sensitive G-protein but also by inhibiting L-type channels, thus reducing the Ca^{2+} signal for exocytosis. The effects on N or P/Q channels are voltage dependent, that is, are evident when the cells are activated by depolarization. These effects last for several seconds after the secretion of enkephalin. In addition, μ-opioids activate large conductance Ca^{2+}-dependent K^+ channels (BK), which reduce Ca^{2+} influx by decreasing the action potential duration when the chromaffin cells are excited by sympathetic nerve stimulation. In this case, there is a tight coupling between the μ-opioid receptor and channel, not involving the G-protein or phosphorylation. Together, these opioid-mediated negative feedback mechanisms (in conjunction with similar ATP-mediated actions) tonically brake the secretory activity of the chromaffin cells. However, with strong stimulation, the voltage-dependent effects of the opioids are lost; thus, the secretion of catecholamines (and enkephalins) is facilitated. The coproduction of enkephalins may thus ensure that there is minimal secretion of catecholamine from the chromaffin cells under basal conditions, but may permit a sharp increase when the animal responds to stress by increasing the frequency of impulses in the sympathetic nerves controlling the adrenal medulla.

In summary, enkephalins cosecreted into the circulation with epinephrine when the adrenal medulla is stimulated can have distant actions. Within the adrenal medulla, enkephalins automatically moderate or inhibit epinephrine release, except when the chromaffin cells are strongly stimulated.

Opioids and Stressor Processing

Pain

The origin of pain is normally at the site of intense stimulation or tissue damage, causing the discharge of impulses in a specialized set of nerve endings (nociceptors). The extent of excitation of these nerve endings is related primarily to the intensity of the painful stimulus, but this is regulated by local chemical signals, including opioid peptides, released when tissue is damaged. Information about painful stimulation is carried as impulses in afferent nerve fibers (especially the small diameter, unmyelinated, and slowly conducting C fibers and the myelinated Aδ fibers, responsible for dull and sharp pain, respectively). These synapse in the dorsal horn, in the substantia gelatinosa (lamina II) and lamina I, respectively, of the spinal cord. Through their long axonal projections in the spinal cord, postsynaptic neurons convey the nociceptive information to the brain. Transmission at these synapses is subject to modulation by locally released neurotransmitters, including opioid peptides. Modulation in the dorsal horn alters the perception of the pain and its impact as a stressor, enhancing (hyperalgesia) or decreasing (analgesia) its effects. The application of morphine to the spinal cord (intrathecal injection or infusion) is highly effective in producing analgesia.

Endogenous opioid and opioid receptor expression in pain-mediating cells There is a high density of all three types of opioid receptors on the nerve fibers in laminae I and II of the dorsal horn. The C (or Aδ) fiber neurons have their cell bodies in the dorsal root ganglia and express μ-opioid receptors in particular, but also δ-receptors, transporting the receptors into both their peripheral and central processes (**Figure 1, iii**). Postsynaptic neurons also have opioid receptors, including κ-opioid receptors. Opioids (enkephalins) in the substantia gelatinosa are produced by the C fiber neurons themselves and by local neurons in the dorsal

horns. Opioids in the dorsal horns include enkephalins, dynorphins (especially in the substantia gelatinosa), and endomorphin 2 (in the nerve fibers of the afferent dorsal root ganglion neurons). Presynaptic opioid actions in the spinal cord involve hyperpolarizing the terminals by activating K^+ channels so that there is less transmitter released when the impulses triggered in the periphery arrive. The main excitatory transmitter released from the C (or Aδ) fiber neurons is glutamate (acting through NMDA receptors), but its effects are enhanced by the coreleased peptide substance P. Substance P and δ-opioid receptors are found together in the secretory vesicles in the nerve terminals of dorsal root ganglion neurons in the dorsal horn. Thus, the opioid receptors are exposed for opioid binding as the nerve terminals release their excitatory neuropeptide. The inhibition by opioids of the release of these excitatory transmitters consequently brakes the activation of relay neurons in the spinal cord that would otherwise have been excited further. Opioids also inhibit the postsynaptic neurons directly, which are probably excitatory interneurons. These oppose the excitatory actions of glutamate through NMDA receptors and, for κ-agonists, also oppose glutamate actions via AMPA receptors.

Opioids in inputs to pain-processing systems The brain influences the processing of nociceptive input in the spinal cord, partly by activating the opioid interneurons in the dorsal horns. There are descending fibers, in the dorsolateral funiculus of the spinal cord, from neurons in the brain stem to the dorsal horns (**Figure 1, i, iii**). These neurons are in the raphe nuclei and locus coeruleus, and they receive input from the PAG and, in turn, from the reticular formation, limbic system, cerebral cortex, and thalamus. They produce monoamines as their transmitters (serotonin, 5-HT; or norepinephrine), and opioids are important in regulating their activity as well as their actions in the spinal cord. When these areas are stimulated, especially the PAG, opioids are released in the dorsal horns, which act via δ-, κ-, or especially μ-opioid receptors and inhibit nociceptive neurons; analgesia results, which is reversed by naloxone. Opioid application in these brain-stem regions has the same effect as stimulation. Hence, endogenous opioids acting on these groups of neurons in the brain, via δ-, κ-, or μ-opioid receptors, are important in controlling or filtering indirectly the flow of nerve impulses from the spinal cord that cause the sensation of pain. Opioids may act directly on these neurons or on the terminals of their afferent neurons; for example, met^5-enkephalin acts like this after its release from presynaptic terminals on locus coeruleus neurons. Appropriately, the descending mechanism is activated by stressors that are likely to lead to pain, giving rise to stress-induced analgesia. Pain pathways continue to ascend to the thalamus and onward to the limbic system and cerebral cortex, affecting mood and bringing the sensation of pain into consciousness. The analgesic actions of morphine include actions at the level of the limbic system and cerebral cortex, which modify further the impact of a painful stressor on the activity of the HPA axis and other stress response mechanisms.

Stress-Induced Analgesia

It has long been recognized that in the heat of battle the perception of pain from an injury can be suppressed. Experimental studies on rats, first reported in 1976, showed that inescapable foot shock increases the pain threshold and that this is partly reversed by naloxone, indicating the involvement of endogenous opioids. The increase in pain threshold is generally tested with a different stimulus from that inducing analgesia, such as, the measurement of tail-flick latency. The mechanism of this stress-induced analgesia involves the pathway previously described, from the brain to the spinal cord; thus, other neurotransmitters involved include monoamines (norepinephrine, 5-HT, and histamine). Microinjections of antagonists have shown that opioid mechanisms (δ- and κ-, as well as μ-, opioid receptors) are activated in the brain stem and spinal cord (**Figure 1, i, iii**). There are differences between individuals or strains of rats or mice in the analgesic response to a stressor, and these are related to genetic factors and the density of μ-opioid receptors in the brain. Also, transgenic mice that have a disruption of the prodynorphin gene, or that do not produce β-endorphin, do not show swim-stress-induced analgesia, which is also suppressed in normal mice by the administration of a κ-opioid receptor antagonist. Stress-induced analgesia can be conditioned, and the conditioned response (analgesia) is prevented by naloxone; thus, the expectation of a painful or stressful stimulus evidently activates brain opioid mechanisms to reduce the anticipated pain. In particular, δ- and μ-opioid receptors (rather than κ-opioid receptors) are involved. An analogous opioid-mediated analgesic effect of anticipated pain has been shown in humans. Such conditioning involves higher brain levels. In contrast, lesion studies show that the PAG and structures more rostral are not necessary for analgesia induced by physical pain (foot shock), which is mediated via the brain stem, particularly the nucleus raphe alatus, and by the spinal cord. Presumably, the ascending information about the foot shock pain stimulates the opioid mechanisms in the brain stem.

Because a stress response (adrenocortical activation) is not essential for stress-induced analgesia,

although basal corticosteroid secretion has a permissive role, it has also been termed adaptive pain suppression. The involvement of opioids in this analgesia varies according to the pattern of stressful stimulation (naloxone reverses analgesia following intermittent but not continuous foot shocks) and its nature (analgesia with tepid- but not cold-water swimming is naloxone reversible and shows morphine cross-tolerance). Clearly, endogenous opioids are important in the analgesia produced by some, generally less persistent and less intense, stressors but not others. In contrast, nociceptin can block supraspinal analgesia, and the potentiation of mild stress-induced analgesia by an ORL-1 receptor antagonist indicates that endogenous nociceptin limits this analgesia. Nonopioid analgesia is a consequence of activation by other transmitters of the descending pathway that opioids can also activate; evidently, endocannabinoids act in the PAG to mediate nonopioid stress-induced analgesia. Lesion of the arcuate nuclei (the location of β-endorphin neurons) reduces opioid-mediated stress-induced analgesia, and transgenic mice with an inactivation of the POMC (β-endorphin) gene do not show normal swim-stress analgesia. In contrast, mice with an inactivation of the preproenkephalin A gene still show opioid-mediated stress-induced analgesia. Thus, β-endorphin, rather than an enkephalin, is the important opioid peptide in the brain.

In mice, social defeat induces analgesia, which is due to endogenous opioid action because it is prevented by naloxone and shows cross-tolerance to morphine. Because adrenalectomy or hypophysectomy does not affect this stress induction of analgesia, it is caused by opioids within the brain rather than by circulating opioids. In contrast, the removal of the pituitary gland or adrenal medulla (sources of β-endorphin and enkephalins, respectively) reduces the analgesia induced by foot shock or immobilization. Evidently, circulating opioids as well as opioids released within the brain can be involved in this analgesia.

An inescapable stressor can evoke learned helplessness (effects caused by the uncontrollability of events that are beyond the organism's control rather than by the events *per se*) as well as analgesia, and this too is opioid-mediated. The inhibition of motor activity, together with a reduced sensitivity to pain, may help the individual to cope or survive. Posttraumatic stress disorder may be modeled by the consequences of exposing an animal to inescapable electric shocks, with opioid-mediated stress-induced analgesia as a common feature.

In summary, pain control by opioids acting within the brain and spinal cord involves automatic feedback control at synapses in pain pathways *pari passu* with noxious stimulation and activation of opioid neurons according to previous experience. The consequence is that these opioid mechanisms contribute to reducing pain sensitivity temporarily, for instance, during acute stress. Morphine mimics these actions to suppress pain.

Opioids and Immune Mechanisms: Inflammatory Pain

The functions of immune cells are modified significantly by the actions of glucocorticoids, and these actions are a major feature of the body's response to stress. However, the chemical messengers, including opioids and CRH, involved in the hypothalamus in regulating ACTH secretion also have important actions directly on immune cells.

Endogenous opioid and opioid receptor expression in immune cells The immune cells themselves produce opioids. T and B lymphocytes, monocytes, and macrophages contain the mRNAs for β-endorphin and met-enkephalin, and produce the peptides; expression is increased at sites of inflammation, and, significantly, immune cells contain pro-protein convertases, which process the opioid precursors. Furthermore, immune cells express the CRH1 receptor and CRH that is released locally from nerve endings stimulates inflammatory responses. One effect of CRH is to stimulate the release, from immune cells, of opioid peptides (especially β-endorphin, a product of the POMC precursor), particularly from lymphocytes migrating to the site of injury; some cytokines act similarly, as does lipopolysaccharide (bacterial endotoxin). The released opioid can then act locally to regulate the inflammatory process (**Figure 1, iv**).

Opioid effects on immune responses Morphine affects both cell-mediated and humoral immune mechanisms. Opiate (heroin) addicts have impaired immune mechanisms. This has implications for the progression of human immunodeficiency virus (HIV) infection in such individuals. Similarly, morphine has immunosuppressive effects in mice. *In vitro* studies showed that opioids (e.g., met-enkephalin acting through μ- or δ-opioid receptors) have direct inhibitory effects on a range of functions of immune cells, including the inhibition of killer activity (NO production) in macrophages and leukocytes. Lymphocytes express δ-, κ-, and μ-opioid receptors, and responses are mediated by $G_{i/o}$-proteins and effects on calcium mobilization. Cytokines are signaling peptides important in the local regulation of immune cell function, and opioids inhibit the cytokine-induced activation and chemotaxis of granulocytes, acting to

desensitize cytokine receptors. However, the rapid effects of opioids can promote immune cell activity, and morphine or opioid peptides (β-endorphin or met-enkephalin) act as chemoattractants to leukocytes. Also, enkephalins stimulate the proliferation of B and T lymphocytes.

In addition, opioids released by immune cells, especially by polymorphonuclear leukocytes, at the site of inflammation have local analgesic actions, hence modulating stressful stimulation. Opioids, including morphine and enkephalins, applied to sites of injury, act through μ-opioid receptors (δ-opioid receptors are also expressed) on the nerve terminals, responding to painful stimuli (nociceptors) and to the chemicals released by local inflammatory cells, and reduce pain (**Figure 1, iv**). These antinociceptive actions are more effective, especially via μ-opioid receptors, if there is also local inflammation. This increased effectiveness is a result of increased opioid receptor expression in the C and Aδ nerve-fiber terminals. It is partly a result of the disruption of the nerve sheath so that opioids have easier access. Opioids act on the nerve terminals to reduce their excitability, probably by opening K^+ channels, hyperpolarizing the terminals. This reduces both the initiation of nerve impulses and the release from these nerve terminals of chemical messengers (e.g., substance P), which promote the inflammatory process. The inhibitory action of opioids on immune cells will reduce their release of chemical pain messengers. Stress, such as swimming in cold water, activates the opioid antinociceptive mechanism in inflamed tissue. This perhaps indicates that circulating opioids, released during stress, may act peripherally at such sites (**Figure 1, iv**). The duration of action of peripherally released opioid peptides is limited by their inactivation by extracellular metallopeptidases, produced by immune cells themselves. Thus, it seems that the stimulated local release of opioids in the initial stages of inflammation may first promote activity of the immune cells and then dampen the response. At the same time, these local opioids reduce inflammatory pain. The nervous and immune systems interact in these processes. The development of drugs to target the peripheral opioid-mediated analgesic mechanisms may afford a new approach to controlling inflammatory pain.

Chronic inflammation leads to changes in the opioid mechanisms that modulate the processing of nociceptive input in the spinal cord. The expression of opioid receptors and opioid peptides in the dorsal horn afferents and neurons shows plasticity. In experimental polyarthritis (a stressful, painful chronic inflammatory condition in which pain sensitivity is locally increased, hyperalgesia), preproenkephalin and prodynorphin mRNAs in the dorsal horn are strongly upregulated and the release of dynorphin in the spinal cord is increased. However, preproenkephalin mRNA expression in dorsal root ganglion neurons is reduced, as is enkephalin release in the spinal cord. Hence, the modulation of pain input in chronic inflammation changes from enkephalin to dynorphin mediation at this level. Nociceptin expression in different dorsal root ganglion cells is increased, and the proposed pronociceptive actions of this peptide may contribute to the hyperalgesia. There is an increase in the number of μ-opioid receptors (and a decrease in δ- and κ-opioid receptors) in the dorsal root ganglion neurons, but this may not result from increased gene expression. Opioid receptor upregulation in the dorsal horn lamina II does involve increased μ- and δ-opioid receptor gene expression in neurons in laminae I and II. These changes may explain the rapid increase in the effectiveness of morphine in inhibiting nociceptive dorsal horn neurons after the induction of focal peripheral inflammation. Clearly, because the chronic inflammatory condition is painful, the changes described do not succeed in preventing such pain.

In summary, the peripheral and central actions of opioids in the brain and spinal cord are complementary in controlling the transmission of information through pain pathways from a site of injury or inflammation. Opioids released at the site of inflammation, or arriving from the anterior pituitary or adrenal medulla, reduce the activity of immune cells and have local analgesic actions. Endogenous opioids reduce the perceived strength of the damaging stressor and thus dampen the HPA axis response to stress; with chronic inflammation, this mechanism may not function.

Opioids and Hedonic Appetite in Stress

It has been known for a long time that opiates and endogenous opioids stimulate the desire for sweets and that this craving is also a feature of stress. Naloxone reduces this desire in humans and in objective tests in laboratory animals. Hence, it has been proposed that the activation of endogenous opioid mechanisms is involved, when an individual is under stress, in enhancing the incentive value of rewarding stimuli, including the pleasant sensations from eating sweet and fatty foods. Consequently, endogenous opioids and their receptors are implicated as a cause of obesity due to hedonically driven overeating in chronically stressed individuals. A consequence of such overeating in animal models is the reduction of the activation of the HPA axis, evidently by a signal from the enlarged abdominal fat store.

Evidence that endogenous opioids are released as a result of consuming sweet foods includes the finding

that this results in analgesia reversible by naltrexone, a long-acting version of naloxone. Studies on mice with transgenic manipulations of the POMC and proenkephalin-A genes so they do not produce β-endorphin or enkephalins showed a reduced incentive to obtain sweet nutrient over and above their need for calories; these findings implicate both of these endogenous opioids in the stimulation of hedonic appetite. In contrast, although nociceptin can also stimulate feeding, it does not stimulate hedonic appetite, so is unlikely to be involved in increasing the desire for sweets and fatty food in stress.

The causal relationship between the activation of brain endogenous opioid mechanisms and the palatability of food or drink extends to the desire to drink alcohol in susceptible individuals; accordingly, naltrexone can be prescribed for the treatment of alcoholism. Further evidence that some actions of nociceptin are functionally opposed to those of opioids is that nociceptin given in the brain in rats acts like naloxone to reduce the stress-induced reinstatement of alcohol-seeking behavior in rats with a previously established preference for alcohol. This has prompted research to develop a nonpeptide ORL-1 agonist to treat alcoholism.

The involvement of endogenous opioids in the normal regulation of appetite and food intake without stress, studied for over 40 years, is complex. Multiple opioid peptides and multiple receptors are involved at several sites in the brain. For instance, male mice genetically engineered so they do not produce β-endorphin eat more and become obese, indicating an anorexigenic role for this opioid in normal appetite regulation and contrasting with its orexigenic actions on hedonic appetite. Because naloxone still reduces food intake in these mice, it is clear that opioids other than β-endorphin usually mediate the orexigenic actions that have previously been suspected to be due to β-endorphin. Sites in the brain where opioids have been shown to stimulate eating have been identified by seeking evidence of opioid modulation of the activity of pools of neurons in discrete brain areas known to be involved in appetite regulation (**Figure 1, v**). These studies involved the microinjection of opioid receptor agonists and antagonists into these brain regions, sometimes into different sites in the same animal, to investigate opioid-mediated interconnections. The effects of such local opioid interventions on appetite, measured in a range of tests with food intake as an end point, and on the activities of neurons in the treated area have then been studied. Neuronal activity has been evaluated principally by the postmortem identification of stimulated immediate-early gene expression, using Fos immunocytochemistry. Of particular interest in the context of stress-induced eating is that fat or sugar intake is stimulated by the microinjection of μ- or δ-, but not κ-, opioid agonists into the shell of the nucleus accumbens. This nucleus receives a dopaminergic input from the ventral tegmental area and is a key part of the mesocorticolimbic system, the reward circuitry that is implicated in drug dependence, including alcoholism. Opioids act in the shell of the nucleus accumbens, via μ- and δ-opioid receptors to increase the release of dopamine, which underlies hedonic sensation. Here the local injection of a μ-opioid receptor agonist stimulates the ingestion of a sweet or fatty meal, and especially a fatty one – increasing almost threefold the amount of fat ingested. The double-injection technique (opioid agonist microinjected into one site, and opioid antagonist microinjected into another) has shown there are opioid-mediated interactions between the nucleus accumbens and the central nucleus of the amygdala, with a μ-opioid agonist stimulating feeding and reciprocal opioid connections between the nucleus accumbens shell and the ventral tegmental area. This reciprocal opioid connection promotes feeding via μ- and δ-opioid receptors. Naltrexone, given systemically, activates neurons in the nucleus accumbens shell, as indicated by Fos expression, and prevents the activation of presumably different neurons in this structure that otherwise follows the ingestion of a palatable meal. Appropriately, the consumption of a sweet, fatty meal activates central β-endorphin neurons. These opioid-driven mechanisms for individuals preferring to ingest energy-rich sweet and fatty foods, notably when they are stressed, can be seen as having a survival advantage in that energy storage is favored. However, with the superabundance of such foodstuffs in advanced societies, these opioid-based reward mechanisms evidently contribute to obesity as a consequence of stress.

The involvement of multiple opioids and their multiple receptors in appetite control, and the multiple other roles of endogenous opioid mechanisms, means that opioid antagonists are insufficiently precise for use in obesity treatment. Instead, recent discoveries have revealed the roles of endocannabinoids and several nonopioid neuropeptides in the regulation of appetite, including α-melanocyte stimulating hormone (MSH), which, like β-endorphin, is a POMC derivative produced by arcuate nucleus neurons, and these have become the focus of attention as novel drug targets.

In summary, endogenous opioids regulate activity in the reward circuitry in the brain, increasing appetite for sweet and fatty foods and the desire for alcohol. The seeking of such hedonic rewards is increased by stress, and opioids drive this behavior.

Opioids, Stress, and Reproduction

Opioids and the Hypothalamic-Pituitary-Gonadal Axis

Reproductive capability depends primarily on the intermittent, approximately hourly, synchronized discharge of gonadotropin-releasing hormone (GnRH) neurons in the anterior hypothalamus/preoptic area. The GnRH so released into the hypothalamo-hypophysial portal system drives gonadotropin secretion from the anterior pituitary gland, which in turn stimulates the sex steroid- and gamete-producing activities of the ovaries or testes. The suppression of this GnRH pulse generator leads to failing gonadal function and to diminishing reproductive capacity or sexual desire. The striking consequence in females of many species is the failure of ovulation and in women the accompanying cessation of the menstrual cycle. Because these effects are a consequence of the suppression of the GnRH pulse generator, this is hypothalamic amenorrhea. Stress depresses reproductive activity in this way, although actions of stress hormones on the anterior pituitary gonadotrophs or gonads may also be involved. Stress inhibits the activity of GnRH neurons through the indirect actions of CRH within the brain via an opioid mechanism in both males and females.

Endogenous opioid effects on gonadotropin secretion during stress: males and females There are many β-endorphin terminals, along with CRH nerve terminals, in contact with, or in the vicinity of, the GnRH neurons (**Figure 1, i**). It is important to note that endogenous opioids also play an important role in the normal regulation of GnRH neurons in active reproductive life, being involved in the interplay between the sex steroid hormones and the neurons. Furthermore, it does not follow that because opioids have been involved in suppressing reproductive function that simply antagonizing opioid actions will reactivate GnRH neurons. Different processes are involved in reactivation. Hence, in women with amenorrhea associated with intensive physical training, naloxone may have no effect. The suppression by restraint stress of the luteinizing hormone (LH) surge in rats that occurs every 4–5 days and triggers ovulation is evidently not caused by endogenous opioids because naloxone does not prevent this.

An acute stressor in intact males can increase LH and testosterone secretion transiently, perhaps aiding aggressive responses. In castrated male rats, foot shock stress decreases LH secretion rapidly, which is reversed by the injection of antiserum to β-endorphin or dynorphin or an irreversible μ-opioid receptor antagonist into the cerebral ventricles, with no effects of δ-opioid receptor ligands. β-Endorphin or CRH injected centrally inhibits LH secretion. Furthermore, CRH is not effective after β-endorphin antiserum, so the inhibition by CRH evidently requires β-endorphin neurons. However, ablation of the PVN, and hence of the parvocellular CRH neurons, does not prevent the foot-shock-induced inhibition of the GnRH pulse generator in male rats and CRH neurons in the PVN evidently do not project to the GnRH neurons, so CRH neurons elsewhere may exert the inhibitory effects of stress on reproductive function. Also, calcitonin gene-related peptide (CGRP), given centrally, suppresses LH pulses, and this action, which is partly via the CRH neurons, is prevented or reversed by naloxone, indicating the mediation of CGRP actions by opioids. In male rats, the inhibition of LH secretion by acute (but not chronic) restraint is reversible by naloxone and hence involves endogenous opioids. Similarly, in female marmosets, aggression by a conspecific causes the naloxone-reversible inhibition of the GnRH pulse generator.

Fasting and insulin-induced hypoglycemia, produced experimentally, are stressors, and they reduce the frequency of GnRH pulses. This reflects a generally inhibitory effect of reduced energy availability on the GnRH pulse generator, so that reproductive activity is suppressed when it would strain further the individual's ability to support its own needs. Leptin is an important signal from adipose tissue in this regard, and the suppression of central kisspeptin production is involved; kisspeptin is a hypothalamic peptide essential for the reactivation of GnRH neurons at puberty. The inhibition of the GnRH pulse generator by fasting or hypoglycemia, which is evident only in females or males with sufficient estrogen or testosterone, respectively, is mediated through μ-opioid receptors; a similar suppression in rhesus monkeys does not involve opioids. The requirement for sex steroids may relate to their stimulatory effects on POMC neurons in the arcuate nucleus. CRH can, with vasopressin, stimulate the release of β-endorphin from these POMC neurons, and this could be the link for the hypoglycemic inhibition of the GnRH pulse generator. Opioids from the adrenal medulla may be involved because the removal of the adrenal medullae abolishes this inhibition. Glucose deprivation stimulates PVN CRH neurons by increasing norepinephrine release, inhibiting GnRH neurons indirectly. The inhibitory actions of CRH on the cell bodies of the GnRH neurons are partly blocked by local μ-opioid receptor antagonists, and thus endogenous opioids released in the vicinity may mediate these CRH actions. However, GnRH neurons do not express

opioid receptor genes, so opioid actions are indirect. Specifically, opioid action in the medial preoptic area seems important, where β-endorphin-containing terminals are found, with noradrenergic and neuropeptide Y (NPY)-containing terminals, on neurons projecting to the PVN. Dynorphin and enkephalin-containing terminals are also found in the vicinity of GnRH neurons. In addition, CRH can act by stimulating endogenous opioid release (β-endorphin or met-enkephalin) in the median eminence, with consequent inhibition, via μ-opioid receptors, at the level of the GnRH nerve terminals. It is not known which endogenous opioid neurons are activated by stress to inhibit GnRH neurons.

Endogenous opioid effects on gonadotropin secretion during stress: males In males, stress decreases the circulating testosterone levels, as we expect from the inhibition of LH secretion. Leydig (interstitial) cells of the testis are the principal source of the male sex steroid hormone, testosterone, but also express the POMC gene, producing β-endorphin under stimulation by gonadotropin along with dynorphins. Sertoli cells (nurse cells for spermatogenesis in the seminiferous tubules) have opioid receptors and produce enkephalins. Opioids affect spermatogenesis through Sertoli cells and testosterone production via any Sertoli cell effects on the nearby Leydig cells. Locally administered naloxone can increase testosterone production. These mechanisms are evidently activated during stress so that immobilization for several hours decreases the circulating level of testosterone by approximately half, due to an action of opioids within the testes involving reduced sensitivity to gonadotropins. It is not clear how the stressor can activate intratesticular opioid mechanisms, but this seems to be a way in which the sex drive may be indirectly reduced during stress.

In summary, endogenous opioids, in particular β-endorphin, mediate the suppression of the pulsatile activity of GnRH neurons and hence mediate the inhibition of gonadotropin secretion and gonadal function that may occur with stress.

Conclusion

In sum, endogenous opioids act as censors, forbidding excessive stimulation by noxious inputs, opposing the actions of chemical messengers acting on neural circuits that favor greater responses to stressors. Endogenous opioids can make the outside world seem less hostile, albeit less sexually interesting, and increase the appetite for rich food.

See Also the Following Articles

Hypothalamic-Pituitary-Adrenal; Anatomy of the HPA Axis; Oxytocin; Pain.

Further Reading

Amit, Z. and Galina, Z. H. (1986). Stress-induced analgesia: adaptive pain suppression. *Physiological Reviews* **66**, 1091–1120.

Appleyard, S. M., Hayward, M., Young, J. I., et al. (2003). A role for the endogenous opioid beta-endorphin in energy homeostasis. *Endocrinology* **144**, 1753–1760.

Bodnar, R. J., Lamonte, N., Israel, Y., et al. (2005). Reciprocal opioid–opioid interactions between the ventral tegmental area and nucleus accumbens regions in mediating mu agonist-induced feeding in rats. *Peptides* **26**, 621–629.

Brunton, P. J., Meddle, S. L., Ma, S., et al. (2005). Endogenous opioids and attenuated hypothalamic-pituitary-adrenal axis responses to immune challenge in pregnant rats. *Journal of Neuroscience* **25**, 5117–5126.

Condren, R. M. and Thakore, J. H. (2001). Cushing's disease and melancholia. *Stress* **4**, 91–119.

Coggeshall, R. E. and Carlton, S. M. (1997). Receptor localization in the mammalian dorsal horn and primary afferent neurons. *Brain Research Reviews* **24**, 28–66.

Dickenson, A. H. (1995). Spinal cord pharmacology of pain. *British Journal of Anaesthesia* **75**, 193–200.

Drolet, G., Dumont, E. C., Gosselin, I., et al. (2001). Role of endogenous opioid system in the regulation of the stress response. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **25**, 729–741.

Fabbri, A. (1990). The role and regulation of testicular opioids. *Trends in Endocrinology and Metabolism* **1**, 117–120.

Fields, H. L. and Basbaum, A. I. (1978). Brainstem control of spinal pain-transmission neurons. *Annual Review of Physiology* **40**, 217–248.

Grossman, A. (1988). Opioids and stress in man. *Journal of Endocrinology* **119**, 377–381.

Levine, A. S. and Billington, C. J. (2004). Opioids as agents of reward-related feeding: a consideration of the evidence. *Physiology & Behavior* **82**, 57–61.

Mansour, A., Fox, C. A., Burke, S., et al. (1994). Mu, delta, and kappa opioid receptor mRNA expression in the rat CNS: an in situ hybridization study. *Journal of Comparative Neurology* **350**, 412–438.

Mercer, M. E. and Holder, M. D. (1997). Food cravings, endogenous opioid peptides, and food intake: a review. *Appetite* **29**, 325–352.

Meunier, J-C. (1997). Nociceptin/orphanin FQ and the opioid receptor-like ORL1 receptor. *European Journal of Pharmacology* **340**, 1–15.

Olszewski, P. K. and Levine, A. S. (2004). Minireview: characterization of influence of central nociceptin/orphanin FQ on consummatory behavior. *Endocrinology* **145**, 2627–2632.

Pasternak, G. W. (2004). Multiple opiate receptors: deja vu all over again. *Neuropharmacology* **47**(supplement 1), 312–323.

Pechnik, R. N. (1993). Effects of opioids on the hypothalamo-pituitary-adrenal axis. *Annual Review of Pharmacology and Toxicology* **33**, 353–382.

Pecoraro, N., Reyes, F., Gomez, F., et al. (2004). Chronic stress promotes palatable feeding, which reduces signs of stress: feedforward and feedback effects of chronic stress. *Endocrinology* **145**, 3754–3762.

Rivier, C. and Rivest, S. (1991). Effect of stress on the activity of the hypothalamic-pituitary-gonadal axis: peripheral and central mechanisms. *Biology of Reproduction* **45**, 523–532.

Russell, J. A., Brown, C. H. and Caron, R. W. (1998). Endogenous opioids. In: Neill, J. D. & Knobil, E. (eds.) *Encyclopedia of reproduction* (vol. 1), pp. 1043–1060. San Diego, CA: Academic Press.

Russell, J. A. and Brunton, P. J. (2005). Neuroactive steroids attenuate oxytocin stress responses in pregnancy. *Neuroscience* **138**, 879–889.

Salzet, M., Vieau, D. and Day, R. (2000). Crosstalk between nervous and immune systems through the animal kingdom: focus on opioids. *Trends in Neuroscience* **23**, 550–555.

Schafer, M., Mousa, S. A. and Stein, C. (1997). Corticotropin-releasing factor in antinociception and inflammation. *European Journal of Pharmacology* **323**, 1–10.

Stamford, J. A. (1995). Descending control of pain. *British Journal of Anaesthesia* **75**, 217–227.

Stefano, G. B., Scharrer, B., Smith, E. M., et al. (1996). Opioid and opiate immunoregulatory processes. *Critical Reviews in Immunology* **16**, 109–144.

Tilbrook, A. J., Turner, A. I. and Clarke, I. J. (2002). Stress and reproduction; central mechanisms and sex differences in non-rodent species. *Stress* **5**, 83–100.

Twitchell, W. A. and Rane, S. G. (1993). Opioid peptide modulation of Ca^{2+}-dependent K^{+} and voltage-activated Ca^{2+} currents in bovine adrenal chromaffin cells. *Neuron* **10**, 701–709.

Yeomans, M. R. and Gray, R. W. (2002). Opioid peptides and the control of human ingestive behaviour. *Neuroscience and Biobehavioral Reviews* **26**, 713–728.

Optimism, Pessimism, and Stress

M F Scheier
Carnegie Mellon University, Pittsburgh, PA USA
C S Carver
University of Miami, Coral Gables, FL USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M F Scheier and C S Carver, volume 3, pp 99–101, © 2000, Elsevier Inc.

Models of Motivation and Optimism
Goals, Stress, and Optimism
Effects of Optimism on Well-Being
Future Directions

Glossary

Coping	Attempts to deal with obstacles to goal attainment and with losses.
Cross-sectional study	A study in which the predictor and outcome variables are assessed at same point in time.
Expectancy-value theories	Models of motivation in which expectancies and values (goals) interact to determine behavior.
Explanatory style	A person's characteristic way of explaining prior negative (and positive) outcomes.
Optimists	People who generally expect to obtain favorable personal outcomes in the future.
Prospective study	A study in which the predictor and outcome variables are assessed at some appropriate baseline and then the outcome variables are reassessed at a later time.
Stress	Experience that arises when people have difficulty moving toward their goals.

Models of Motivation and Optimism

Optimists are people who hold generalized positive expectations for the future. They expect good things to happen to them, not bad things. Pessimists are just the reverse. They expect to miss out on the good things in life and to experience things that are bad. Research suggests that individual differences in optimism and pessimism may play an important role in the manner in which people react to stressful circumstances. To understand the nature of the effects, it is important to understand the broader theoretical context in which research on optimism and pessimism is

embedded. It is also important to understand how goals and the pursuit of goals are involved in creating stress.

Expectancies are pivotal in most contemporary theories of optimism. Defining optimism and pessimism in terms of expectancies creates a link to expectancy–value theories of motivation (some of which date back to the first few decades of the twentieth century). Expectancy–value models assume that behavior occurs in order to attain desired values or goals. If engagement of effort is to occur, there must be a goal that matters enough (has enough value) to try to reach it. The other element in the equation is expectancy: Confidence or doubt that the goal will be obtained. If the person lacks confidence, again there will be no action. If the person loses confidence along the way, action will stop. Only when confidence is sufficiently high will the person act and remain engaged in goal-directed efforts.

Expectancies exist at many levels of generality. A patient lying in a hospital bed recovering from an automobile accident can have a variety of different expectancies. She can have an expectancy about being able to move her index finger one more centimeter forward, an expectancy about being able to get out of bed unaided by next Saturday, and an expectancy about fully recovering from the trauma in the long run. Presumably, the principles embodied by expectancy–value theories pertain equally well to expectancies that are specific in nature and expectancies that are more general. They should even apply to the most general kinds of expectancies, those that characterize optimists and pessimists. The confidence that is at issue is simply broader in scope, relevant to a greater class of situations and behaviors.

Goals, Stress, and Optimism

Expectancies are important facets of expectancy–value theories, but so are goals and values. The emphasis on goals meshes well with aspects of contemporary theorizing in personality and social psychology. Recent years have seen a reemergence of interest in goal constructs. One prevailing view is that human behavior is organized around the pursuit of goals – trying to attain things that are desired and avoid things that are not. Because of differences in emphasis, theorists use different terms to refer to goals, including current concerns, personal strivings, personal projects, and life tasks. Goals that are even broader in nature carry labels such as possible selves and self-guides.

Although these various goal constructs certainly differ among themselves in ways that are far from trivial, they also share an underlying commonality. All include the idea that goals energize and direct activities. All include the sense that goals give meaning to people's lives and that understanding the person means understanding the person's goals and the organization of those goals. As such, goals are seen as providing the structure that define people's lives, imbuing lives with meaning, both in the short run and the long run.

Goal constructs provide an interesting window on the experience of stress and coping. From this vantage point, stress occurs when people experience difficulty moving toward desired end states or difficulty keeping away from end states that are unwanted. Stated somewhat differently, stress occurs whenever impediments to goal attainment are encountered. From this perspective, coping involves efforts to create conditions that permit a person to continue to move in the right direction (toward the desired goals or away from the undesired goals) or to disengage from goals that are seen as no longer attainable. Thus, stress and coping are not seen as distinct classes of phenomena but, rather, are seen as arising within the dynamics of normal goal-related self-regulatory processes.

What determines how a person reacts to the blockage of goal-directed activities? Expectancy–value theories suggest that the key determinant is the person's confidence or doubt that goal attainment will eventually occur. Thus, when confronting minor challenges, optimists take a posture of confidence and continue to persist in their goal-directed efforts, even if progress is difficult or slow. Pessimists are more doubtful and reserved in their efforts. This divergence should also be displayed as the threat to goal attainment becomes more extreme. Optimists are likely to believe that the adversity can be handled successfully; pessimists are likely to anticipate disaster. To the extent that continued engagement in goal-directed efforts is adaptive, optimists should experience a coping advantage over pessimists.

Effects of Optimism on Well-Being

Measuring Optimism

Research on the beneficial effects of optimism has flourished over the past 20 years. This research has tended to take two different routes to assessing optimism, leading to two somewhat distinct literatures. One approach measures expectancies directly, asking people to respond to statements such as "I'm optimistic about my future." The second approach is more indirect. It assesses expectancies by examining attributional style, the characteristic manner in which a

person explains prior events. Bad outcomes that are explained in terms of causes that persist into the future, influence a broad range of events, and involve an aspect of the self are thought to be pessimistic in nature in that they carry the implication that negative outcomes will continue to occur. The opposite attributional style – explaining negative events in terms of causes that are more time limited, narrow in their effects, and external to the self – are thought to reflect a more optimistic orientation. Although in principle attributional style can be assessed with respect to both negative and positive prior events, in practice assessment is usually limited to negative outcomes only.

Psychological Well-Being

Dozens of studies have been conducted examining the relationships between optimism, pessimism, and distress among groups of people undergoing adversity of one type or another. The range of stressors embodied by this research has been extremely varied. Studies have examined the experiences of students entering college, employees of businesses, and survivors of missile attacks. Studies have measured the reactions of people caring for cancer patients and people caring for patients suffering from Alzheimer's disease. Research has examined the experiences of people dealing with medical procedures such as childbirth, abortion, coronary artery bypass surgery, and attempts at *in vitro* fertilization, as well as heart and bone marrow transplantation. Yet other studies have looked at how people deal with a diagnosis of cancer, the pain of arthritis, and the progression of AIDS.

The results of these various studies all point in the same direction: optimists experience less distress during times of adversity than do pessimists. This conclusion holds for studies that are cross-sectional in nature – that is, those that assessed optimism and distress at the same point in time. This conclusion also holds true for studies that are prospective in nature – that is, those that assessed optimism and distress at some appropriate baseline and then reassessed distress at a later time. The results of the prospective studies are particularly noteworthy in that they suggest that optimism is associated with beneficial changes in distress over time. Thus, these prospective studies also help to get around the potential problem of confounding optimism with subjective well-being, which is inherent in the cross-sectional research.

Physical Well-Being

Much more is known about the effects of optimism on psychological well-being than is known about the effects of optimism on physical well-being. Still, a number of studies have explored the links between optimism and physical health and between optimism and parameters of physiological functioning. The findings from this body of research mirror those just presented with respect to psychological functioning: optimists typically show signs of better physical health or signs of more adaptive physiological responses when under adversity than do pessimists.

For example, compared to pessimists, optimists report fewer physiological symptoms during times of duress and maintain a higher health status across their lives. Other studies show that optimists are less likely to suffer negative side-effects from major surgery or to be rehospitalized within the first few months after major surgery. In a similar vein, less pessimistic people have been found to outlive people who are more pessimistic in outlook when diagnosed with a life-threatening illness such as recurrent cancer. They also exhibit less extreme cardiovascular reactivity during the course of their daily lives. On a somewhat different note, optimists tend to show signs of more adaptive immune functioning than do pessimists, but the evidence on this point is more mixed than on the other physical-health outcomes that have been examined.

Future Directions

At present, the evidence strongly indicates that optimism confers important benefits. Still, a number of questions remain for future research to resolve. First, there is a question concerning the structure of optimism and pessimism. Researchers tend to treat optimism and pessimism as though they were opposite poles of a single continuum. However, some evidence suggests that optimism and pessimism may represent related but independent dimensions. This is one question that researchers have been trying to sort out for a while and will continue to sort out into the future.

Second, little research has been directed toward understanding the developmental antecedents of optimism and pessimism. We do know from several studies that optimism and pessimism have a genetic basis. Less is known about the environmental influences on optimism and pessimism. What accounts for the nongenetic variation in optimism and pessimism? What kinds of life experiences cause a person to become more or less optimistic? These will also be important questions for future research to address.

Third, what can be done to turn pessimists into optimists? To date, relatively little research has concerned itself with generating and evaluating interventions designed to reduce pessimism. Given the toxic effects

of a pessimistic orientation, the future will no doubt see a significant increase in the number of studies that are designed to develop effective interventions to counteract the negative effects of pessimism.

Finally, it now seems well established that coping differences underlie the differences in psychological adjustment that emerge in the manner in which optimists and pessimists adjust to stress. Less clear are the biological and physiological mechanisms that underlie the differences in physical health outcomes that have been observed. The search for the biological pathways by which optimism gets translated into health outcomes will also occupy a substantial portion of the research landscape in the future.

Acknowledgments

The preparation of this article was facilitated by NIH grants HL065111, HL065112, HL076858, HL076852, CA64710, and CA84944.

See Also the Following Article

Personality Processes.

Further Reading

Carver, C. S. and Scheier, M. F. (1998). *On the self-regulation of behavior*. New York: Cambridge University Press.

Carver, C. S. and Scheier, M. F. (1999). Stress, coping, and self-regulatory processes. In: Pervin, L. A. & John, O. P. (eds.) *Handbook of personality* (2nd edn., pp. 553–575). New York: Guilford Press.

Carver, C. S. and Scheier, M. F. (2002). Optimism. In: Snyder, C. R. & Lopez, S. J. (eds.) *Handbook of positive psychology*, pp. 231–243. New York: Oxford University Press.

Carver, C. S. and Scheier, M. F. (2003). A self-regulatory perspective on personality. In: Millon, T. & Lerner, M. J. (eds.) *Comprehensive handbook of psychology. Vol. 5: Personality and social psychology*, pp. 185–208. New York: John Wiley.

Pervin, L. (ed.) (1989). *The goal concept in personality and social psychology*. Hillsdale, NJ: Lawrence Erlbaum.

Peterson, C. and Bossio, L. M. (1991). *Health and optimism: new research on the relationship between positive thinking and well-being*. New York: Free Press.

Scheier, M. F. and Carver, C. S. (2003). Goals and confidence as self-regulatory elements underlying health and illness behavior. In: Cameron, L. D. & Leventhal, H. (eds.) *The self-regulation of health and illness behavior*, pp. 17–41. Reading, UK: Harwood.

Scheier, M. F., Carver, C. S. and Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a re-evaluation of the Life Orientation Test. *Journal of Personality and Social Psychology* **67**, 1063–1078.

Scheier, M. F., Carver, C. S. and Bridges, M. W. (2001). Optimism, pessimism, and psychological well-being. In: Chang, E. C. (ed.) *Optimism and pessimism: implications for theory, research, and practice*, pp. 189–216. Washington, DC: American Psychological Association.

Scheier, M. F., Matthews, K. A., Owens, J. F., et al. (1999). Optimism and rehospitalization following coronary artery bypass graft surgery. *Archives of Internal Medicine* **159**, 829–835.

Seligman, M. E. P. (1991). *Learned optimism*. New York: Knopf.

Orexin

W K Samson and M M White
St. Louis University, St. Louis, MO, USA
A V Ferguson
Queen's University, Kingston, ON, Canada

© 2007 Elsevier Inc. All rights reserved.

Discovery of the Orexins
Orexins and the Stress Response
Orexin and Stress-Related Behaviors
Summary and Conclusion

Glossary

Adrenocorticotropic hormone (ACTH)	A stress hormone that activates hormone secretion from the adrenal gland.
Cerebroventricles	Fluid-filled cavities in the brain where test compounds often are injected *in vivo*.
Circumventricular organs	Sites in the brain where there is no blood–brain barrier.
Corticosterone	A steroid hormone released from the adrenal gland during stress that facilitates energy delivery to the nervous system and stimulates the production of new glucose.

Hypothalamus	A brain region responsible for coordinating the appropriate endocrine, cardiovascular, and behavioral responses to stress.
Prolactin (PRL)	A hormone released from the anterior pituitary gland in stress that exerts multiple tissue-specific effects.
Rostral ventrolateral medulla	An important site in the hindbrain where neurons organize and control sympathetic efferent activity.
Stress hormones	Chemical messengers released from the anterior pituitary gland during physical or emotional stimuli.

Discovery of the Orexins

The lateral hypothalamic area (LHA) has long been considered to be an important site for the regulation of feeding behavior. In 1998, two groups independently identified the production of novel peptides in this region. deLecea and colleagues named these peptides hypocretins because of their apparent production in the hypothalamus and their structural homology to the members of the secretin family of peptides. Sakurai and collaborators named these peptides orexins on the basis of their ability when administered intracebroventricularly (ICV) to stimulate food intake. It is now recognized that the peptides exert many physiologically relevant actions in brain, the majority of which are seemingly related to behavioral, endocrine, and cardiovascular responses to stress.

For convenience we use here the nomenclature advanced by Sakurai. Orexin A and orexin B are postranslational products of the same prohormone (Figure 1). Two receptor proteins have been identified, the orexin type 1 receptor (OX_1R), which binds both peptides but prefers orexin A as a ligand, and the orexin type 2 receptor (OX_2R), which binds both peptides with approximately equal affinity. The two receptors are differentially distributed, as listed in Table 1.

The distribution of the orexin receptors predicted more biological actions than simply those related to food intake. Indeed, receptor localization in the dorsal raphe, tuberomammillary nuclei, and locus ceruleus (LC), areas known to be important in arousal state and sleep/wakefulness, predicted the eventual finding that the orexins are important determinants of sleep/wakefulness. The absence of the receptors in dogs and degeneration of orexin-expressing neurons in animal models and humans results in the syndrome narcolepsy/cataplexy. Thus, the major role recognized at this time for these peptides is one of arousal. Aspects of the narcolepsy/cataplexy syndrome highlight other reported biological actions of the orexins, particularly those related to stress.

Orexins and the Stress Response

Orexins and Cardiovascular Function

Soon after their discovery, the orexins were demonstrated by several groups to act in the brain to elevate

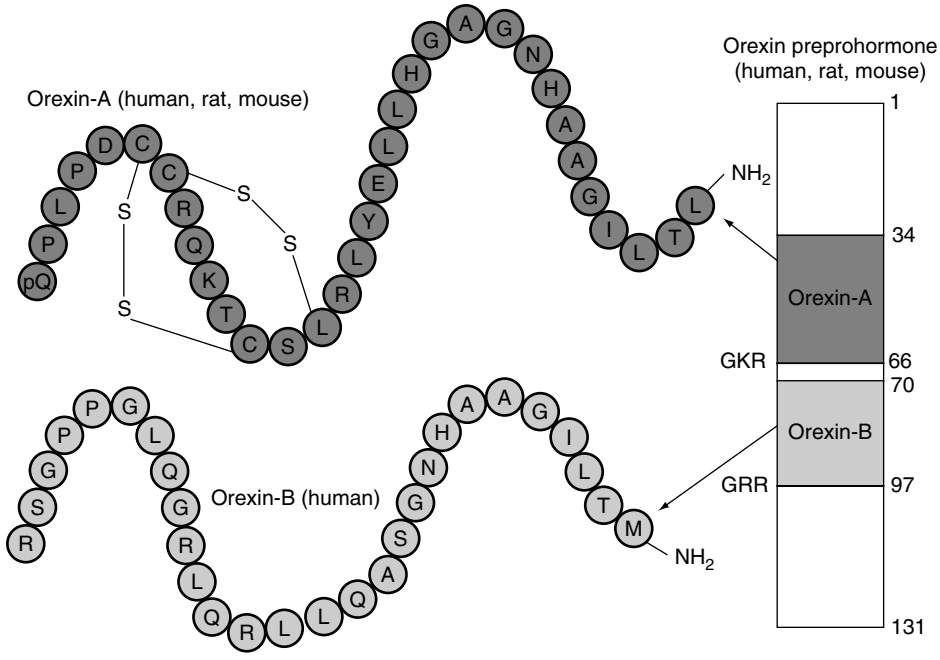


Figure 1 Structure of the orexins. From *Trends in Endocrinology and Metabolism* volume 11, 2000.

Table 1 Distribution of orexin receptor subtypes in brain regions important to the stress response[a]

Brain site	*OX_1R mRNA*	*OX_2R mRNA*
Bed nucleus, stria terminalis	High	Moderate
Medial amygdala (nucleus)	Moderate	Low
Cortical amygdala (nucleus)	Low	Moderate
Periventricular thalamus (nucleus)	High	Moderate
Medial preoptic hypothalamus (area)	Absent	Moderate
Medial preoptic hypothalamus (nucleus)	Moderate	Moderate
Organum vasculosum lamina terminalis	Low	Moderate
Median preoptic area	Low	Moderate
Anterior hypothalamic nucleus	High	Low
Hypothalamic arcuate nucleus	Absent	High
Lateral hypothalamus (area)	Low	High
Medial paraventricular nucleus (parvocellular)	Absent	High
Lateral paraventricular nucleus (parvocellular)	Absent	Low
Posterior paraventricular nucleus (magnocellular)	Absent	Absent
Tuberomammillary nuclei	Absent	High
Ventromedial hypothalamus (nucleus)	High	Moderate
Dorsomedial hypothalamus (nucleus)	Absent / low	Moderate
Midbrain A4, A5, and A7 cell groups	High	Absent
Dorsal motor nucleus of vagus	Moderate	Moderate
Dorsal raphe nucleus	Moderate	Moderate
Locus coeruleus	High	Low
Nucleus of the solitary tract	Moderate	Low
Parabrachial nucleus	Absent	Low
Periaqueductal grey	Moderate	Moderate
Sensory and spinal trigeminal nuclei	Absent	Moderate
Substantia nigra (pars compacta)	Moderate	Moderate
Ventral tegmental area	Moderate	Moderate
Zona incerta	Moderate	Moderate

[a]As determined by *in situ* hybridization histochemistry. Density of the histochemical signal detected is indicated as high, moderate, low, or absent (at background).

blood pressure and heart rate. The effects are mediated via the activation of the sympathetic nervous system, as demonstrated by the use of α-adrenergic antagonist blockade (phentolamine) and the demonstration of increased sympathetic nerve activity. Increased cardiovascular function was observed following the ICV or direct tissue-specific administration of the peptides into the area postrema, rostral ventrolateral medulla, and medial nucleus tractus solitarius; the direct effects on preganglionic sympathetic neurons in the intermediolateral cell column (IMDL); and the ability of intrathecal administration of orexin to activate these neurons. In addition, orexin has been shown to influence the excitability of paraventricular nucleus (PVN) parvocellular and nucleus tractus solitarius (NTS) neurons. This does not rule out the potential contribution of vagal withdrawal to the observed increased heart rate following central orexin administration. In fact, direct effects of orexin on neurons in the dorsal motor nucleus of the vagus also have been reported.

Cellular mechanisms of action In an attempt to determine the cellular mechanism of action of orexin to raise blood pressure and heart rate by an action in the NTS, experiments were conducted in medullary slice preparations. Whole-cell recordings from NTS neurons in current clamp mode revealed that the depolarizing effect of orexin was maintained in the presence of tetrodotoxin, demonstrating direct postsynaptic actions of the peptide. In voltage clamp mode, orexin was observed to inhibit a specific potassium conductance (the sustained K^+ current) and to activate a nonselective cation conductance, leading to depolarization. Later a G-protein signaling mechanism was demonstrated to activate phospholipase C and protein kinase C but not protein kinase A (PKA).

The cellular effects of orexin also have been characterized in the area postrema (AP), a circumventricular organ located adjacent to the NTS. In dissociated AP neurons, orexin similarly activated a nonselective cation conductance, leading to depolarization. In these cells, however, no effects on potassium currents were observed. The activation of a nonselective cation conductance also has been demonstrated in parvocellular PVN neurons, which may in fact have been preautonomic cells projecting

to sympathetic centers in the rostral ventrolateral medulla. As in the AP, there did not appear to be any action of orexin on potassium currents in neurons in parvocellular PVN.

In sympathetic preganglionic neurons in spinal cord, orexin's depolarizing effects are blocked by pertussis toxin, indicating a G-protein signaling mechanism. In these neurons, a PKA-mediated inhibition of potassium channels has been demonstrated. The direct effects of orexin to depolarize noradrenergic neurons in the LC occur due to decreased potassium conductance and the activation of an inward current. The exact receptor subtype responsible for these depolarizing actions of orexin in the NTS, AP, PVN, LC, or IMDL has not been identified; however, the similar (and in some cases diverse) signaling mechanisms described predict heterogeneity of receptor subtypes depending on tissue localization.

Physiological relevance of the pharmacological effects of orexin on cardiovascular function Mice in which the gene for orexin has been deleted (embryonic knockouts) and mice and rats engineered to suffer age-related loss of orexin neurons (orexin-ataxin 3 transgenes) display lower mean arterial blood pressures than their wild-type littermates under resting conditions and fail to mount a sympathetic response (increased heart rate or increased blood pressure) in the resident/intruder stress model. The reason for the lower resting blood pressures in these transgenic animals was established to be due to a loss of vasoconstrictor sympathetic tone because prazocin (α-adrenergic blockade) or hexamethonium (ganglionic blockade) failed to alter the already low pressures in these animals but brought pressures in wild-type littermates to levels similar to the transgenic animals. This was not accomplished by treatment with captopril (angiotensin-converting enzyme inhibitor) or a V_1 (vasopressin antagonist).

Does the syndrome narcolepsy/cataplexy in humans demonstrate compromised sympathetic function? Autonomic dysfunction has been reported in these patients, including the attenuation of cardiovascular reflexes to muscle contraction and the Valsalva maneuver. Frequent occurrences of orthostatic hypotension have been reported as well.

Neuroendocrine Actions of the Orexins

Orexin and the hypothlamic-pituitary-adrenal axis The presence of orexin receptors (Table 1) and broad distribution of orexin-positive nerve fibers in the hypothalamus suggested actions of the peptides on neuroendocrine centers in the diencephalon or the pituitary gland itself. Initial attention was drawn to potential actions in anterior pituitary gland because orexin receptors are abundant in the gland and the peptide appears to be delivered to the median eminence, where it could gain access to the hypophysial portal circulation. Although no significant effect of the orexins on prolactin (PRL) release were observed in dispersed anterior pituitary cell cultures, the release of the other major anterior lobe stress hormone, adrenocorticotropic hormone (ACTH), was significantly affected by the concentrations of peptide in the nanomolar range. Basal ACTH release was not affected by orexin; however, corticotropin releasing hormone (CRH)-stimulated ACTH release was inhibited in a concentration-related fashion. The effect appeared to be mediated by the OX_1R because the median effective concentration (EC_{50}) for orexin A was at least 1 log lower than that for orexin B and because the inhibitory action of orexin was absent when the production of the OX_1R (but not the OX_2R) was compromised by antisense oligonucleotide treatment. Recently, the selective OX_1R antagonist has been demonstrated to also block the action of orexin A on CRH-induced ACTH release *in vitro*. Furthermore, the relatively selective OX_2R agonist (Ala^{11}, Leu^{14}) orexin B did not exert actions similar to orexin A or orexin B. Orexin's action on the corticotroph appear to require the activation of a non-pertussis toxin-(non-PTX-) sensitive G-protein-coupled receptor to phospholipase C activation and the formation of inositol trisphosphate. In addition, the activation of PKC seems to be a necessary step in transduction of the inhibitory signal.

In the hypothalamus, orexin acts to stimulate the hypothalamic-pituitary-adrenal (HPA) axis by activating CRH-producing neurons. This was demonstrated by several groups using pretreatment with CRH antiserum or antagonists prior to the ICV administration of orexin. Whether this is a direct or indirect action of orexin on CRH-producing neurons in the parvocellular PVN is currently unknown, although single-cell the reverse transcription polymerase chain reaction (RT-PCR) approaches now employed in the authors' laboratories promise a future definitive answer to this question. Neurons in the PVN express both receptor subtypes, although the OX_2R appears more abundant. In addition, orexin-producing neurons that project to the PVN express glucocorticoid receptors, and thus the closed feedback loop of the HPA axis appears to include orexinergic neurons in the LHA (direct effects of orexin, of unknown origin, on adrenal steroid production and release have also been reported). As previously mentioned, direct postsynaptic effects of orexin on PVN neurons have been reported. The orexin effects on parvocellular PVN neurons, those projecting to median eminence or other brain sites (e.g., sympathetic

centers in the brain stem) are direct (i.e., postsynaptic) and are mediated by the activation of nonselective cation conductances. In magnocellular PVN, the depolarizing effects of orexin are indirect (i.e., tetrodotoxin blockable), involving a glutamatergic interneuron. Because high concentrations of arginine vasopressin (AVP) are capable of stimulating ACTH release from cultured pituitary cells *in vitro* and under at least some *in vivo* conditions the compromise of AVP action reduces stress-induced ACTH release, it is possible that the ACTH response to the ICV administration of orexin may be mediated by AVP release into the median eminence. The central administration of orexin has been reported to increase both CRH and AVP mRNA levels in the PVN; however, in hypothalamic explant cultures, orexin did not stimulate AVP release. Thus, it appears that the hypothalamic action of orexin to stimulate the HPA axis is mediated primarily by CRH release into the median eminence.

Orexin and prolactin secretion Controversy exists over the potential action of orexin on PRL secretion. Orexin, in log molar concentrations ranging from 1.0 pM to 100 nM, did not alter basal or thyrotropin-releasing hormone (TRH)-stimulated PRL release from dispersed anterior pituitary cells *in vitro*. However, when it was administered ICV some groups observed stimulatory and others inhibitory effects on PRL levels in plasma. The suppressive effects of orexin on *in vivo* PRL secretion were reported to be the result of increased dopamine turnover in male, but not female, rats. Evidence for a stimulatory role for orexin on PRL and luteinizing hormone (LH) releases come from fasting studies in which the preovulatory PRL and LH surges, absent during fasting, were restored by central orexin administration. Even more support for the stimulatory actions of orexins comes from steroid-priming studies in ovariectomized rats, in which the PRL and LH surges were blocked by pretreatment with anti-orexin antibodies. Whether any of these pharmacological effects of orexin on PRL or LH secretion are physiologically relevant is unclear and may not be likely because orexin knockout mice are fertile and can nurse their pups. It is possible that physiologically relevant actions of the orexins on at least PRL secretion may not be related to reproductive function but instead to stress-induced hormone release. Thus, the actions of the peptide on tuberoinfundibular dopamine neurons of the arcuate nucleus, where the OX_2R is abundantly expressed, may be observed only in models of acute stress such as restraint (physical) or novel environment (anticipatory). The expression of early response genes in neurons of the LHA has been described in several experimental models of stress.

Orexin and Stress-Related Behaviors

The presence of orexin receptors and orexin-positive nerve fibers in central nervous system (CNS) sites known to be important in the arousal state, such as the LC and raphe nuclei, forms the anatomical basis for the effects of pharmacological doses of orexin on motivated and spontaneous behaviors. Several groups have described effects of orexin on locomotor behavior and general arousal state, and the obligatory role for endogenous orexin in sleep/wakefulness is well established. In our hands, the effects of orexin on general locomotor behavior and stereotypic behaviors parallel in dose relationship and time course the actions on sympathetic activity (i.e., blood pressure elevation). In addition, some of the feeding responses to orexin may be related to the peptide's ability to stimulate arousal, as well as exploratory and reward behaviors. These behaviors, observed following central orexin administration, mirror behaviors observed during stress; however, they differ from those observed following CRH administration, which are more anxiogenic in nature. There may be an important interplay between orexin neurons in the LHA and CRH neurons in the PVN. As already detailed, orexin appears to activate CRH neurons (resulting in ACTH release), and it has been reported that CRH receptors are present on orexin neurons in the LHA, where CRH itself exerts depolarizing effects. Thus, at least some of the behavioral arousal stimulated by CRH may be mediated by the effects of the peptide on endogenous orexin release.

The pharmacological actions of orexin on locomotor activity and arousal state are mirrored by determinations of orexin neuronal activity in these behaviors. Indeed, c-FOS expression correlates directly with wakefulness and locomotor activity, and the spontaneous activity of these neurons during such states has been characterized as well. One site of orexin action related to spontaneous locomotor activity is the PVN. In addition to stimulating feeding in the early light phase, orexin injection into the PVN increased spontaneous locomotor activity at all the times examined, and this was accompanied by increased oxygen consumption. The effect of orexin on spontaneous locomotor activity was antagonized by pretreatment with SB-334867 (an OX_1R antagonist). Those results suggest that physical activity independent of feeding is due to the actions of orexin in the hypothalamus. Indeed in recent studies, we observed that many, but not all, of the behavioral actions of exogenous orexin could be antagonized by pretreatment with the OX_1R antagonist, whereas some of the similar behavioral actions of the selective OX_2R agonist remain unaffected by OX_1R

antagonist pretreatment. Thus, both receptors may be responsible for differing behavioral modalities elicited by orexin administration. It will be important to determine the receptor subtype selectivity of the individual behaviors so that the roles played by each receptor in not only locomotor but also stress-related behaviors can be elucidated.

Non-Central Nervous System Sites of Orexin Action Related to Stress

Orexin receptors have been identified in non-CNS tissues, and significantly, the adrenal gland expresses both receptors but the OX_2R more abundantly. Orexin acts both *in vitro* and *in vivo* to stimulate corticosterone release, with orexin A being more potent than orexin B. This suggests a role for the OX_1R in the control of adrenal cortical function. The effect is selective (aldosterone secretion was unaffected) and is blocked by H-89 pretreatment *in vitro*, suggesting a signaling pathway involving adenylyl cyclase. Because the major site of orexin production was the LHA and plasma levels of the peptide are extremely low, it remained unclear that these pharmacological effects of orexin in the adrenal gland had any physiological correlate. However, it is now known that the orexins are also produced in the adrenal gland and thus may exert autocrine or paracrine actions in the gland. It should be remembered that the direction of blood flow in the adrenal gland is from the outer cortex to the inner medulla, and thus orexin produced in the adrenal cortex might bathe the medulla in relatively high concentrations, similar to the situation with corticosterone. At least in human pheochromocytomas, orexins act via the OX_2R to stimulate epinephrine and norepinephrine release, suggesting an action of orexin to regulate catecholamine release from the normal adrenal gland. Whether orexin production and/or release from cortex increases during stress is at this point unknown; however, it is tempting to speculate that the adrenal actions of orexin complement the peptide's actions in brain related to metabolism and stress.

Summary and Conclusion

Abundant evidence support a role for endogenous orexins in the regulation of sleep/wakefulness and metabolic state. Central to all the actions of the orexins is the consistent finding that multiple indicators of stress are revealed following orexin administration and some are absent when orexin production or action is compromised. Remarkable progress has been made in understanding the importance of the endogenous orexin system, largely due to the development of transgenic models of orexin peptide or receptor compromise. In particular, the development of the orexin-ataxin-3 transgenic mice and rats has been a landmark in our ability to understand the broad impact of these peptides on multiple behavioral, endocrine, and cardiovascular systems. Now, with the availability of selective OX_1R antagonists and at least a relatively selective OX_2R agonist, additional elucidation of the sites and mechanisms of orexin actions in the brain can be accomplished. Because the replacement of the peptide restores normal sleep/wakefulness patterns in transgenic animals, the use of orexin or orexin analogs for the treatment of human narcolepsy/cataplexy holds much promise. Critical to the development of those potential therapeutic strategies is the recognition of the receptor subtypes responsible for the effects of orexin on cardiovascular and neuroendocrine function. Clearly, potential side effects related to cardiovascular control and hormone release are possible and should be monitored if orexin is to be tested for its therapeutic effect in humans.

See Also the Following Articles

Food Intake and Stress, Human; Feeding Circuitry (and Neurochemistry).

Further Reading

de Lecea, L., Kilduff, T. S., Peyron, C., et al. (1998). The hypocretins: hypothalamus-specific peptides with neuroexcitatory activity. *Proceedings of the National Academy of Sciences USA* **95**, 322–327.

de Lecea, L. and Sutcliffe, J. G. (eds.) (2005). *Hypocretins: integrators of physiological functions*. New York: Springer.

Ferguson, A. V. and Samson, W. K. (2003). The orexin/hypocretin system: a critical regulator of neuroendocrine and autonomic function. *Frontiers in Neuroendocrinology* **24**, 141–150.

Hara, J., Beuckmann, C. T., Nambu, T., et al. (2001). Genetic ablation of orexin neurons in mice results in narcolepsy, hypophagia, and obesity. *Neuron* **30**, 345–354.

Kayaba, Y., Nakamura, A., Kasuya, Y., et al. (2003). Attenuated defense response and low blood pressure in orexin knockout mice. *American Journal of Physiology* **285**, R581–R593.

Kilduff, T. S. (2005). Hypocretin/orexin: maintenance of wakefulness and a multiplicity of other roles. *Sleep Medicine Reviews* **9**, 227–230.

Kiwaki, K., Kotz, C. M., Wang, C., et al. (2004). Orexin A (hypocretin 1) injected into hypothalamic paraventricular nucleus and spontaneous activity in acts. *American Journal of Physiology* **286**, E551–E559.

Marcus, J. N., Aschkenasi, C. J., Lee, C. E., et al. (2001). Differential expression of orexin receptors 1 and 2 in rat brain. *Journal of Comparative Neurology* **435**, 6–25.

Mieda, M., Willie, J. T., Hara, J., et al. (2004). Orexin peptides prevent cataplexy and improve wakefulness in an orexin-ablated model of narcolepsy in mice. *Proceedings of the National Academy of Sciences USA* **101**, 4649–4654.

Mileykoviskiy, B. Y., Kiyaschenko, L. I. and Siegel, J. M. (2005). Behavioral correlates of activity in identified hypocretin/orexin neurons. *Neuron* **46**, 787–798.

Peyron, C., Tighe, D. K., van den Pol, A. N., et al. (1998). Neurons containing hypocretin (orexin) project to multiple neuronal systems. *Journal of Neurosciences* **18**, 9996–10015.

Sakurai, T., Amemiya, A., Ishii, A., et al. (1998). Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* **92**, 573–585.

Samson, W. K., Taylor, M. M. and Ferguson, A. V. (2005). Non-sleep effects of hypocretin/orexin. *Sleep Medicine Reviews* **9**, 243–252.

Willie, J. T., Chemelli, R. M., Sinton, C. M., et al. (2001). To eat or sleep?: orexin in the regulation of feeding and wakefulness. *Annual Reviews of Neuroscience* **24**, 429–458.

Organ Transplantation, Immune Response *See:* Organ Transplantation, Stress of.

Organ Transplantation, Stress of

M A Dew, A F DiMartini and R L Kormos
University of Pittsburgh School of Medicine and Medical Center, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Prevalence of Organ Transplantation
The Transplantation Process for the Individual Patient
Interventions to Minimize Stressor Effects during the Transplant Process

Glossary

Graft rejection	The process by which the body's immune response system attacks the foreign organ or tissue.
Immunosuppression	The suppression of the immune response by an agent such as a chemical or drug in order to enhance the body's acceptance of foreign organs or tissue.
Organ transplantation	The process of removing an organ or tissue from one person's body and implanting it in another person's body.

Organ transplantation is increasingly often the procedure of choice for patients with end-stage diseases of the kidney, liver, heart, lung, pancreas, or intestines. Organ transplantation can significantly extend life, as well as dramatically improve the quality of individuals' lives. Yet each step in the transplantation process – encompassing the initial evaluation and wait for an organ, the transplant surgery, postsurgical recovery, and living permanently with the transplant – is also associated with significant stressors. These stressors include both acute events and chronic, ongoing health and psychosocial strains. Across the entire process, transplant patients face and must adapt to changing health and functional capacity, altered social relationships, new perceptions of self, and revised life goals and plans. An understanding of both the nature of such stressors and individuals' reactions to them is critical for three reasons. First, as transplantation becomes more common, it is imperative that patients and their families know what to expect during the transplant process and thereafter. Education regarding likely stressors as well as potential benefits of transplantation will increase the likelihood that organ recipients will achieve the best possible quality of life (QOL) in the months and years after transplant. Second, it is only through knowledge about the nature of transplant-related stressors that we will be able to develop and test appropriate stress reduction interventions for transplant recipients and their families. Third, knowledge regarding transplant-related stressors, and reactions to them, also has value for understanding similar reactions to the experience of other life-threatening chronic physical illnesses, their treatment, and their long-term outcomes.

Prevalence of Organ Transplantation

Although the potential benefits of solid organ transplantation were first realized in the 1950s, it was not until the advent of modern immunosuppression in the 1970s and 1980s that organ transplantation fully evolved from an experimental procedure to the standard of care for many end-stage diseases. Today, the greatest obstacle to receiving a transplant is the shortage of donor organs. As illustrated in **Figure 1,** the total number of individuals registered on the wait list for a transplant (including kidney, liver, heart, lung, heart-lung, pancreas, or intestines) far exceeds the number of transplants performed, and the gap continues to widen, despite the success of some social and governmental programs in increasing rates of donation. The lives of kidney transplant candidates can be maintained by dialysis during the wait for an organ. But for other solid organ candidates, for whom few or no such long-term alternative treatments exist, the consequence of the organ shortage is that a large proportion of them will die before a suitable organ becomes available.

In some areas of transplantation, living donor organs have helped to reduce waiting times and allow more candidates to receive transplants. For example, in the United States, living kidney donation – whether from a donor who is biologically or emotionally related to the recipient or from an anonymous, unrelated donor – is almost as prevalent as deceased donor transplantation. Living liver donation is also becoming a viable option for some patients. **Figure 2** shows the total numbers of deceased donor transplants (and living donor transplants, where applicable) that were performed for each type of organ transplant in 2004 in the United States. Overall, kidney transplantation (whether from deceased or living donors) is the most common type of transplant performed, followed by liver, heart, pancreas, and lung transplantation. Intestine and heart-lung transplants are the most rare. Similar distributions are reported in other countries.

As shown in **Figure 3,** survival rates following organ transplantation vary across organ type, but are good to excellent for many types of organs, especially in the early years after transplant. Kidney recipients, especially those with living donors, show the highest rates of survival. Even at 10 years posttransplant, 76% remain alive. Liver and heart recipients show slightly lower survival rates, but, even so, almost 75% survive at least 5 years posttransplant, and about 50% remain alive at 10 years posttransplant. Survival rates for intestine and lung recipients remain substantially poorer, especially in the long-term. But even these rates have shown marked improvements in recent years. Given that annual mortality rates range up to 30% or more among candidates awaiting transplant, it is clear that transplantation provides a significant extension of life for many patients.

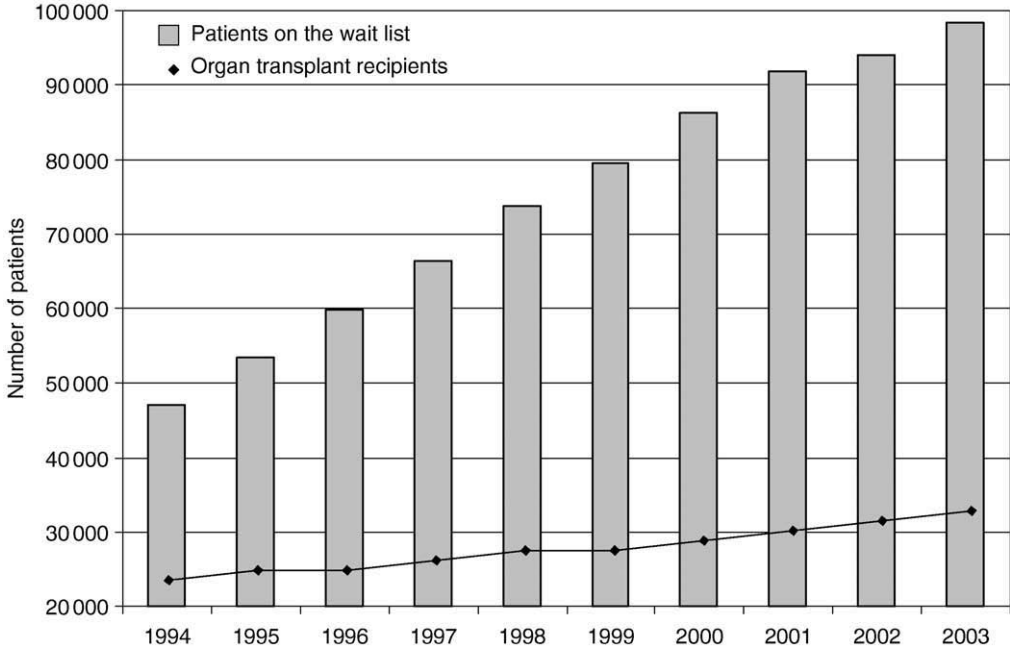


Figure 1 Combined United States and Eurotransplant statistics. From 2004 Annual Reports from United Network for Organ Sharing, Organ Procurement and Transplantation Network, which includes all United States data (www.optn.org), and Eurotransplant, which includes data from Austria, Belgium, Germany, Luxembourg, Netherlands, and Slovenia (www.eurotransplant.org).

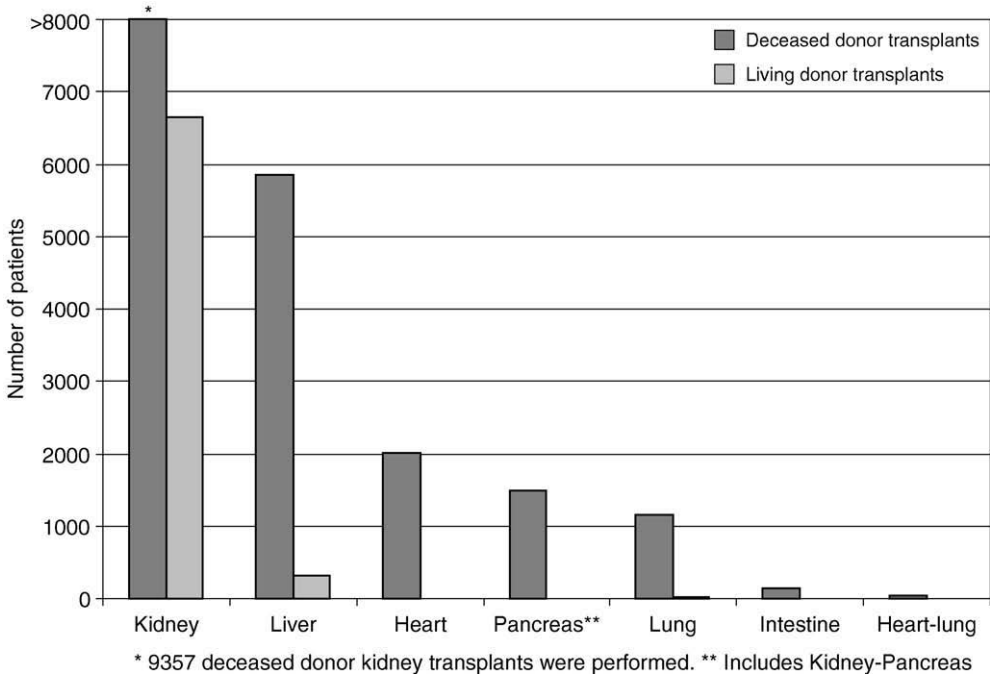


Figure 2 Numbers of deceased donor and living donor transplants of each organ type (including kidney-pancreas) in the United States in 2004; 9357 deceased donor kidney transplants were performed. From United Network for Organ Sharing, Organ Procurement and Transplantation Network (www.optn.org).

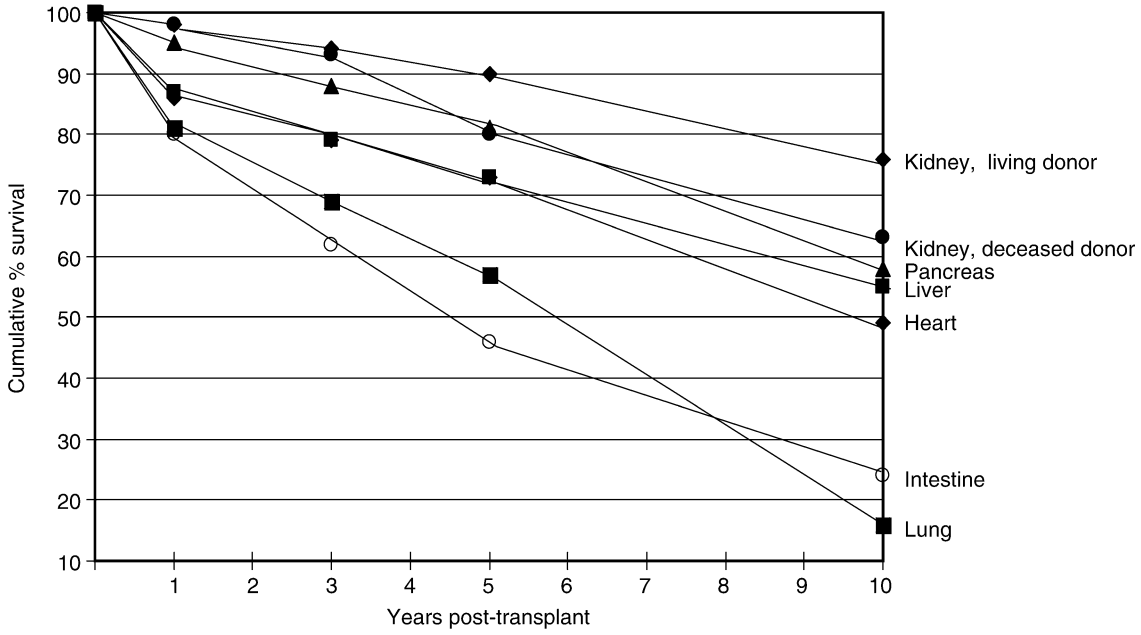


Figure 3 Patient survival, U.S. transplants, 1996–2003. From 2004 Annual Report, United Network for Organ Sharing, Organ Procurement and Transplantation Network.

The Transplantation Process for the Individual Patient

From the point of their initial contact and evaluation by the transplant team, patients and their families move through a relatively standard series of events and time periods associated with the transplant process. The typical time line is shown in Figure 4. The time line is punctuated by specific events that occur during the transplant process, and these events mark the onset and offset of important periods or stages, including the wait for a donor organ, the

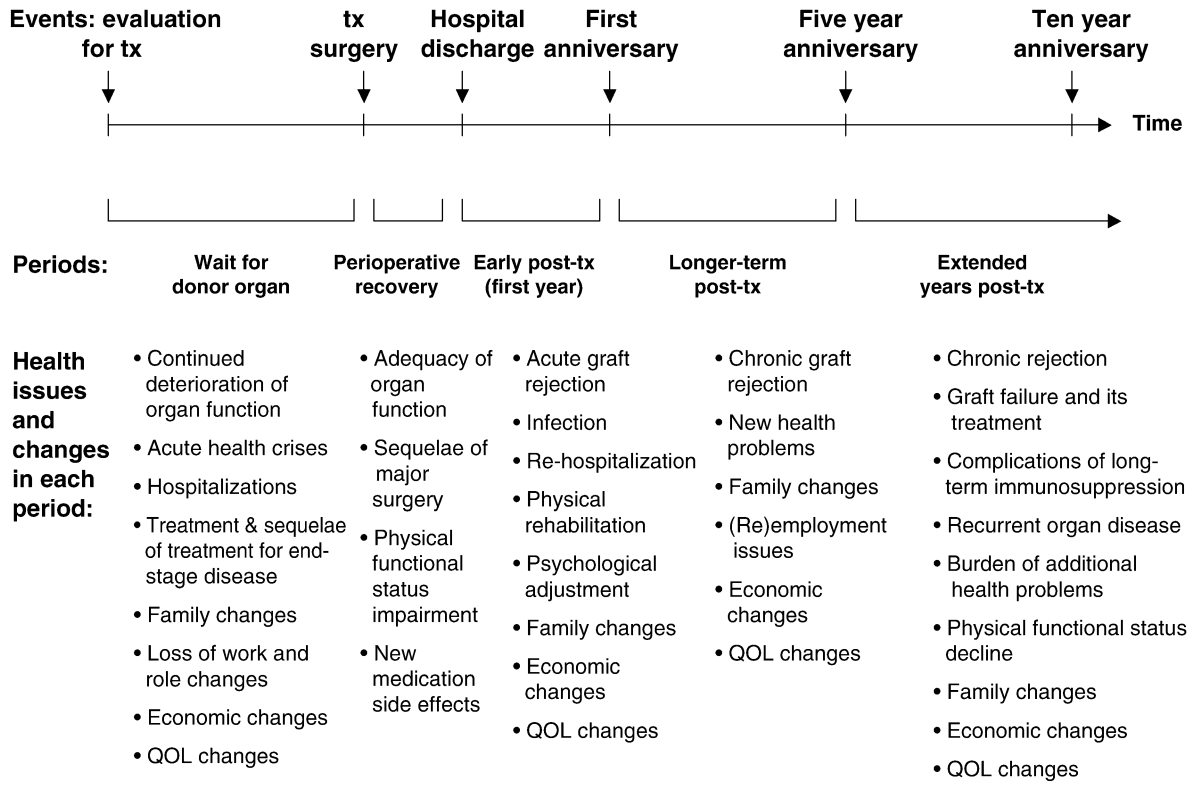


Figure 4 Organ transplant timeline: critical events, time periods, and health and psychosocial issues. Abbreviations: tx, transplantation; QOL, quality of life. Adapted from Dew, M. A. et al. (2002). Psychosocial aspects of transplantation. Reprinted with permission from Medscape Transplantation. © 2002, WebMD, Inc.

perioperative recovery period, the early posttransplant months, and the longer-term posttransplant years. The duration of each of these periods can vary considerably, depending on such factors as the rapidity with which end-stage organ disease develops in a given patient, the availability of a living donor (when living donation is possible) or a suitable deceased donor organ, and the medical complications that may occur perioperatively or later posttransplant. However, the essential ordering of the events and time periods is similar for all patients.

Figure 4 also lists the typical health and psychosocial issues that are associated with each time period. Some of these issues are unique to specific time periods (e.g., continued deterioration of the native organ's function during the waiting period, or early posttransplant physical rehabilitation). Others are present across almost all of the periods (e.g., changes in QOL), although the central elements of concern to patients and their families in these areas vary across time. Each of the sections below considers the stressors and related health and psychosocial issues that are relevant to the major events and the time periods in Figure 4.

The Initial Evaluation for Transplant

The formal medical evaluation for transplant candidacy is often the first occasion during which patients and their families seriously consider transplantation as a therapeutic option. For both patients and their family members, the evaluation may evoke conflicting feelings, including relief and the hope of having a healthy future, fears about surgical risks and adapting to life with an organ from another person, worry about whether the patient will be judged to be a suitable transplant candidate, worry about the financial costs of transplantation, concerns that an organ will not become available soon enough to save the patient's life, and eagerness to try a new treatment option combined with reservations as to whether there might still exist other seemingly less drastic medical interventions. For all of these reasons, studies of psychological distress among patients and family members conducted at the time of the evaluation typically find that many respondents' distress levels are high. Symptoms of anxiety and depression have been noted most frequently, and distress in these areas may reach the level of diagnosable disorder in some individuals.

The evaluation for transplant is itself usually lengthy and complex because it is designed to assess all aspects of patients' medical and psychosocial status. Patients may be hospitalized during the several days of the evaluation in order to complete all required tests and procedures. An extensive array of medical and laboratory tests is required, and virtually all transplant programs consider patients' psychosocial and psychological histories. Although there is considerable variation across programs in the depth with which psychosocial history is probed, important elements that are usually evaluated to at least some degree include lifetime history and treatment of psychiatric disorder; lifetime history and current use of alcohol, nicotine, and other substances; current and past level of adherence to medical regimens; cognitive status and ability to understand what will be required for posttransplant care; social history including employment, financial circumstances, marital status, and availability of social supports from family and friends; personal views and expectations about the prospect of receiving a transplant; and personal strategies for coping with functional limitations due to health.

These psychosocial components are routinely assessed because patients' history and current status in these areas may influence their psychological adaptation to the wait for the transplant and to life after the transplant. However, these assessments often serve as important additional stressors to patients and families. They are stressful because of the personal nature of the questions as well as because the results may be used for decision making by the transplant team. Typically, these evaluations will contribute to judgments as to whether patients should be immediately listed as transplant candidates or whether they need to undergo additional interventions to improve their health status and adherence to the medical requirements set by the transplant team. For example, patients may be required to enroll in alcohol or smoking cessation programs before the transplant team will accept them onto the waiting list.

The Waiting Period

Many patients and their families perceive the waiting period to be the most psychologically stressful part of the transplant experience, due largely to patients' continued physical health deterioration and the inherent uncertainties about whether and when a suitable donor organ will become available. The duration of this period can vary from a few days (although brief waits are now rare) to several years, depending on the patient's medical status and other factors such as blood type and the patient's physical size. The geographical location of the transplant center can be a major determinant of waiting time as well, since organs are allocated in many countries through a regional system that depends on the location from which the donor organ originated. Given the uncertainties of obtaining a donor organ, plus the risks of the transplant surgery itself (especially as the patient's medical condition deteriorates), patients and their families are faced with the mutually opposing prospects of preparing to live and preparing to die.

In the face of these stressors, the risk of significant psychological distress and cognitive impairment rises. Thus, as for many chronic disease populations, transplant candidates are much more likely to experience elevated distress and diagnosable depressive and anxiety-related disorders than nonpatient, community-based populations. Whether a particular transplant candidate is at greater risk for depression or anxiety appears to depend in part on their end-stage organ disease. For example, depression appears to be more common in kidney, liver, and heart candidates, while anxiety disorders – panic disorder in particular – are more often observed in patients with end-stage lung disease.

In general, the longer a patient waits for a transplant, the greater the psychological toll on the patient and family. However, it is noteworthy that an extremely brief wait (e.g., in the case of patients who have very rapidly declining organ function and are fortuitously able to undergo transplantation very quickly) may have negative psychological consequences as well. For example, it has been found that levels of anxiety and rates of anxiety disorders in the first year after transplantation are higher in patients who wait for very brief periods of a few months or less, compared to those who wait for longer periods. It is thus possible that the waiting period – despite its central elements of uncertainty and its associated risk for psychological distress when the wait is prolonged – can provide important time for patients to adapt to the idea that they need a transplant. If they do not have time before the surgery to adjust to the prospect of transplantation, the immediate consequence may be greatly heightened distress as they attempt to adapt to the transplant experience in the early postsurgical period.

Neurocognitive impairments are common in transplant candidates. For example, hepatic encephalopathy is prevalent in liver candidates and ranges from mild cognitive deficits to delirium and coma. Even patients with chronic liver disease who do not have hepatic encephalopathy may have cognitive and behavioral abnormalities as results of more subtle derangements of metabolic and detoxification functions. Cognitive deficits in heart candidates are extremely common once cardiac ejection fraction (a measure of the heart's ability to pump) drops below

30%; when brain perfusion is substantially inadequate, delirium results. Although kidney patients' conditions are generally more stable compared to candidates for other types of solid organs, subtle cognitive deficits often occur when patients are not well dialyzed.

Continued elevations in psychological distress and/or neurocognitive changes may have their own negative effects on other areas of patient well-being. For example, impaired patients may have more difficulty adhering to the medical regimen requirements set by the transplant team: they may have trouble following prescribed diets or abstaining from nicotine, alcohol, or other substances. These difficulties may arise not only due to impaired motivation (e.g., as a result of depression) but as a result of confusion and memory loss.

Finally, as patients' organ function declines, they may require increasingly complicated medical technologies and treatment regimens in order to maintain their lives. These treatments also serve as stressors that further tax patients' strained emotional and cognitive resources. The strain may extend to the family as well. The patient's marriage and relationships with primary family caregivers (who are most often spouses of adult patients and parents of pediatric patients) are particularly vulnerable due to role changes within the family and changes in daily living activities and schedules. The financial ramifications of these changes – arising from medical expenses, the inability of the patient to work, and/or the need for key family members to take substantial time off from work – can further add to strain within the family.

The Surgery and the Perioperative Recovery Period

This phase of the transplant process is characterized by major, but often positive, physical and emotional transitions. With improvements in procedures and medical care in the immediate days posttransplant, hospital stays have been dramatically reduced. For example, patients who receive the most common types of transplant (kidney, liver, or heart) are frequently able to be discharged in 7 to 10 days. However, as with all major surgery, the possibility for medical complications exists, and this, plus the patient's physical status immediately prior to surgery, will affect the speed of perioperative recovery.

During the first few weeks after surgery, most patients and their families voice high levels of optimism about the patient's prospects for recovery. Their central concerns usually involve how well the new organ is functioning and the probable speed of the patient's physical rehabilitation. These concerns are sometimes heightened by the brief postsurgical hospital stays. Moreover, the short duration of the hospitalization is often perceived as stressful because patients and families typically receive many educational materials about life posttransplant during this time. (This information may have been provided before the transplant as well, but patients may have had difficulty absorbing it due to other stressors and/or cognitive impairment.) The positive feelings and optimism that often dominate after the surgery can interfere with educational efforts during the perioperative period because these feelings may lead patients and their families to be less able to accept and focus on information concerning potential complications, financial issues, and family difficulties that can arise posttransplant.

An additional source of stress arises from medication changes made in conjunction with the transplant. The immunosuppressants that transplant recipients will take for the remainder of their lives are introduced with the transplant, and they bring an extensive range of side effects. While some side effects abate with time, transplant patients – just like any other chronic disease population – will find that others appear to be permanent and unpleasant (e.g., tremor, altered body appearance, changes in sexual functioning, appetite changes). The new medications can also precipitate mood changes. For example, the introduction of corticosteroids can precipitate the onset or recurrences of mood or anxiety disorders. In addition, it is not uncommon for patients who were stabilized on psychotropic medications before transplant to have those medications abruptly discontinued after surgery. This can lead to rapid destabilization for patients, whose mood and/or anxiety disorders may quickly recur.

The First Year after the Transplant

For most patients, the first year posttransplant is a period of readjustment and rehabilitation, with gradual improvement in all domains of QOL, including physical, emotional, and social well-being. A frequent source of frustration for both patients and their families is the fact that the recovery process is generally considerably slower than they had expected. Moreover, important psychosocial elements of the transplant process may further prolong recovery and may also affect family members' well-being. These elements pertain to patients' need to (1) cope with common complications such as acute graft rejection episodes and infections, (2) alter their self-image from that of a critically ill or dying patient and resume a less illness-focused lifestyle, and (3) psychologically accept the fact that they have an organ from another person. Aside from these personal issues, patients may feel daunted by continuing financial issues related to the costs of the transplant, follow-up care,

and medications. In addition, in terms of interpersonal relationships, both patients and family members can be dismayed that the same marital, parent–child, or other family-related difficulties that existed before (and/or during) the period of critical illness and transplant surgery continue after the patient returns home. In fact, the stress of the waiting period may have held some relationships together temporarily, and these may then dramatically decline after the transplant.

An additional area of difficulty for many patients is adherence to the complex posttransplant treatment regimen. Not only must patients take multiple medications, but they also must complete routine medical follow-up evaluations and laboratory tests, monitor vital signs at home (e.g., blood pressure, temperature), follow exercise and dietary requirements, and adhere to restrictions on smoking, alcohol, and other substance use. Just as for most chronic disease patients who must adhere to a complex regimen, transplant recipients' adherence tends to worsen with time. This worsening itself can be a source of frustration (as well as a health risk) for patients, who may become quite discouraged at their inability to refrain from long-standing habits (e.g., smoking) or maintain new ones (e.g., exercise).

A likely result of these many concerns is that the likelihood of experiencing high levels of emotional distress, as well as clinically significant episodes of depressive and anxiety-related disorders, is higher during the first year posttransplant than during subsequent years. For most patients, distress levels gradually abate over the course of the year, and some psychiatric disorders – for example, posttraumatic stress disorder (PTSD) related to the transplant – appear to rarely if ever recur in later years. Individuals with PTSD related to the transplant appear to be unable to come to terms with the transplant experience and may experience flashbacks, nightmares, and extreme distress when thinking about one or more of the stressors that occurred during the waiting period, the surgery, and/or the perioperative recovery period.

Not only can the elements that serve as posttransplant stressors exacerbate psychological distress, but also such distress can itself serve as a major source of strain that, in turn, contributes to the worsening of these other stressors. Thus, for example, episodes of depression and anxiety can lead to increased difficulties in adhering to the medical regimen, as well as affect the patient's ability to cope with new medical complications or the need for rehospitalizations. There is evidence that, perhaps via these behavioral mechanisms, psychological distress can lead to poorer clinical outcomes and reduced survival rates after transplant.

The Longer-Term Years after the Transplant

By the end of the first year after transplant, most patients have achieved their maximal level of physical rehabilitation and have psychologically adjusted to the transplant experience. Most then embark on a multiyear period in which graft functioning remains high and the rate of serious or life-threatening medical complications is low. Acute graft rejection episodes may be rare or relatively infrequent, compared to the first year after transplant. Patients' QOL in the areas of physical, emotional, and social well-being is often high, and rates of rehospitalization for rejection, infections, etc., are low. Some patients will develop new health problems, including diabetes, hypertension, and kidney disease, often as a result of long-term immunosuppression use. Patients are also at greater risk for cancers due to these drugs. Generally, however, patients' level of functioning remains quite high during these years.

Focal issues for patients and their families during these years pertain to maintaining as high a level of QOL as possible and postponing the effects of any decline in graft function or other complications for as long as possible. Economic and financial concerns related to costs of medications and continuing health care can become more important as the years go by. Nevertheless, psychological distress related to the transplant experience and its medical or psychosocial sequelae is uncommon during these years. Instead, patients' mental health is much more likely to be influenced by the same sorts of acute life events and chronic life strains experienced by other community populations. Indeed, the fact that nontransplant life stressors assume increasing prominence as predictors of emotional difficulties during these years attests to the fact that the vast majority of patients – even those who had great psychological difficulty immediately posttransplant – eventually are able to incorporate the experience into their lives and go on to focus on other things.

The Very Extended Years after the Transplant

We know the least about the stressors and psychosocial outcomes experienced by transplant recipients and their families beyond the 5-year anniversary of the transplant. Even in the 1980s and early 1990s, survival rates were sufficiently poor beyond this point that most clinical and psychosocial research did not consider these later-term years. Currently, however, with large proportions of patients surviving even well

Table 1 Examples of interventions that have been evaluated for their ability to improve patient and family psychosocial outcomes during the organ transplantation process

Intervention	*Sample and study groups*	*Design*	*Key results*	*Reference*
Education classes (3x/week for 6 weeks) plus exercise and weight training sessions (2x/week for 6 weeks), compared to education alone. Education focused on lung disease, home health care, stress reduction techniques	Nine lung transplant candidates randomized to education plus exercise (n = 5) or education alone (n = 4)	Randomized controlled trial; assessments pre- and postintervention, no blinding	No significant between-group changes, but global QOL and physical functional status significantly improved in both groups. No other improvements in physical, emotional, or social well-being	Manzetti, J. D., Hoffman, L. A., Sereika, S. M., et al. (1994). Exercise, education, and quality of life in lung transplant candidates. *Journal of Heart and Lung Transplantation* **13**, 297–305
Telephone-based sessions (1x/week for 8 weeks) compared to usual care (medical management by transplant team). Telephone sessions were supportive counseling with cognitive-behavioral techniques to address stress, health, and coping	71 lung transplant candidates randomized to telephone (n = 36) or usual care (n = 35)	Randomized controlled trial; assessments pre- and postintervention; transplant team blinded to treatment assignment; assessors not blinded	Telephone group had lower somatic, anxiety, and depression symptoms and better global QOL, global mental health, and perceived social support postintervention, controlling for baseline levels. No group differences in physical functional QOL or perceived stress	Napolitano, M. A., Babyak, M. A., Palmer, S., et al. (2002). Effects of a telephone-based psychosocial intervention for patients awaiting lung transplantation. *Chest* **122**, 1176–1184
Individual compared to group psychotherapy (1x/week for 12 weeks). Each condition used systemic integrative psychotherapy, which focuses on patient understanding of problems and means to address them	89 kidney recipients randomized to individual (n = 49) or group therapy (n = 40). A control group receiving no treatment (n = 37) was also enrolled	Partially randomized trial, with the control group enrolled after all other subjects; assessment pre- and postintervention (or at similar intervals in controls), with 3-, 6-, 9-, and 12-month follow-up. No blinding	Both intervention groups showed significant reductions in depressive symptoms that were maintained through 1 year postintervention. Control group showed considerably lower depressive symptoms than intervention groups at baseline, and symptom levels rose over time. By the final assessment, symptom levels in the three groups were similar	Baines, L. S., Joseph, J. T. and Jindal, R. M. (2004). Prospective randomized study of individual and group psychotherapy versus controls in recipients of renal transplants. *Kidney International* **65**, 1937–1942
Psychosocial intervention (4 months exposure to Internet-based program) compared to usual care (medical management by transplant team). Internet intervention included components to address health education, stress management, psychological well-being, and medical compliance	60 heart recipients receiving either the intervention (n = 20) or usual care (n = 40). Each recipient's primary family caregiver was also enrolled	Controlled trial with historical comparison group (usual care); assessments pre- and postintervention (or at similar intervals in controls); transplant team blinded to intervention participation; assessors not blinded to intervention participation but blinded to study hypotheses	Intervention group patients had significantly improved depression and anxiety symptoms, and caregivers had significantly improved anxiety and anger symptoms, relative to the comparison group. Social functioning QOL significantly improved in intervention patients and caregivers (not assessed in comparison group). Patient medical compliance did not change, but the subgroup who used the component of the intervention focused on the medical regimen showed significant improvement in keeping clinic appointments, completing blood work, and following diet	Dew, M. A., Goycoolea, J. M., Harris, R. C., et al. (2004). An internet-based intervention to improve psychosocial outcomes in heart transplant recipients and family caregivers: development and evaluation. *Journal of Heart and Lung Transplantation,* **23,** 745–758

Mindfulness-based stress reduction program (1x/week for 8 weeks). Program consisted of training sessions and daily practice of meditation and gentle yoga techniques	20 transplant recipients (2 lung, 12 kidney, 5 kidney-pancreas, 1 pancreas)	Single-group trial; assessments pre- and postintervention and at 3 months follow-up; no blinding	Sleep, depressive symptoms, and overall mental health significantly improved pre- to postintervention. At 3-month follow-up, sleep improvement was maintained, but depressive symptoms had worsened. At 3-month follow-up, anxiety symptoms were improved from baseline. No changes in global QOL, physical functional QOL, or medical adherence	Gross, C. R., Kreitzer, M. J., Russas, V. et al. (2004). Mindfulness meditation to reduce symptoms after organ transplant: a pilot study. *Advances in Mind-body Medicine* **20**, 29–29
Telephone-based QOL therapy compared to telephone-based support therapy (1x/week for 8 to 12 weeks). QOL therapy used cognitive-behavioral strategies to address patient-identified problems in various QOL domains. Supportive therapy involved listening and providing information about the transplant experience	35 lung transplant candidates randomized to QOL therapy ($n = 17$) or supportive therapy ($n = 18$)	Randomized controlled trial; assessments pre- and 1-month, 3-month postintervention; transplant team blinded to treatment assignment	While the groups were equivalent at baseline, the QOL therapy group showed significantly better global QOL, lower mood disturbance, and a better relationship with spouse/partner at 1 month postintervention. QOL and mood improvements were maintained at 3-month postintervention	Rodrigue, J. R., Baz, M. A., Widows, M. R, et al. (2005). A randomized evaluation of quality-of-life therapy with patients awaiting lung transplantation. *American Journal of Transplantation* **5**, 2425–2432

beyond 10 years posttransplant, it is critical that the stressors, as well as the continued benefits, of transplantation be more carefully delineated.

Many of the complications that began but remained at low levels of severity in earlier years become symptomatic and can lead to major functional limitations beyond 5 years posttransplant. For example, chronic graft rejection, kidney disease that progresses to kidney failure, complications of diabetes, and cancers all increase in prevalence. In all types of organ transplant recipients, kidney failure, due to long-term immunosuppressant use, can precipitate the need for a kidney transplant. Many individuals who had originally received nonrenal transplants (i.e., liver, heart, lung) do not expect to eventually need another type of transplant, and learning that they now need a kidney transplant can reawaken the depression and anxiety that they experienced while awaiting their earlier transplants. Individuals who originally received kidney transplants can experience a profound sense of bereavement as their renal graft fails and they return to dialysis.

Recipients of nonrenal transplants may develop chronic rejection that is sufficiently severe that they require a retransplant with a new heart, liver, or pancreas. Retransplantation of nonrenal organs remains considerably less successful than first-time transplantation. Therefore, progressive graft failure for these individuals is likely to be associated with increasing psychological distress because treatment alternatives and the potential for good outcomes are poor.

In short, the accumulating medical burden experienced by all types of transplant recipients in the extended years posttransplant is likely to lead to decrements in all areas of well-being and QOL. Some of the health-related stressors and their likely sequelae will be similar to those experienced during the end-stage organ disease that necessitated the original transplant. Other stressors may be more similar to those experienced by other chronic disease and late-life populations as they approach the upper end of their life expectancies.

Interventions to Minimize Stressor Effects during the Transplant Process

Transplant programs have attempted a variety of strategies to help patients and their families cope with transplant-related stressors and/or to minimize the stressors' impact on patients' emotional well-being and QOL. These interventions have included educational, psychotherapeutic, and self-help support groups, patient-to-patient mentoring programs, and the provision of educational and referral materials using a variety of audiovisual formats. The effectiveness of the majority of these programs has not been empirically evaluated. A very small literature to date has tested psychosocial interventions that could be employed either face to face when organ candidates and recipients return to the medical center for care or by telephone or via the Internet. Examples of these interventions and study findings regarding improved patient and family psychosocial outcomes are listed in Table 1. More extensive intervention development and testing for effectiveness are critical needs in this field.

Acknowledgments

This article was supported in part by grant MH072718 from the National Institute of Mental Health, Rockville, MD, USA.

See Also the Following Articles

Disease, Stress Induced; Psychosomatic Medicine.

Further Reading

Cupples, S. A. and Ohler, L. (eds.) (2002). *Solid organ transplantation: a handbook for primary health care providers*. New York: Springer Publishing Co.

Dew, M. A., Switzer, G. E., DiMartini, A. F., et al. (2000). Psychosocial assessments and outcomes in organ transplantation. *Progress in Transplantation* **10**, 239–259.

Dew, M. A., Manzetti, J., Goycoolea, J. M., et al. (2002). Psychosocial aspects of transplantation. In: Smith, S. L. & Ohler, L. (eds.) *Organ transplantation: concepts, issues, practice and outcomes*, chap. 8. New York: Medscape Transplantation, WebMD, Inc. (www.medscape.com/transplantationhome).

DiMartini, A. F., Dew, M. A. and Trzepacz, P. T. (2005). Organ transplantation. In: Levenson, J. L. (ed.) *The American Psychiatric Publishing textbook of psychosomatic medicine*, pp. 675–700. Washington, D.C.: The American Psychiatric Press, Inc.

Olbrisch, M. E., Benedict, S. M., Ashe, K. and Levenson, J. L. (2002). Psychological assessment and care of organ transplant patients. *Journal of Consulting and Clinical Psychology* **70**, 771–783.

Rodrigue, J. R. (ed.) (2001). *Biopsychosocial perspectives on transplantation*. New York: Kluwer.

Trzepacz, P. T. and DiMartini, A. F. (eds.) (2000). *The transplant patient: biological, psychiatric and ethical issues in organ transplantation*. Cambridge, UK: Cambridge University Press.

Oxidative Stress

H Sies
Heinrich-Heine-Universität, Düsseldorf, Germany
D Jones
Emory University, Atlanta, GA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by H Sies, volume 3, pp 102–104, © 2000, Elsevier Inc.

Oxidative Stress Definitions
Related Terms
Reductive Stress
Nitrosative Stress
Concluding Remarks

Glossary

Oxidative stress	An imbalance between oxidants and antioxidants in favor of the oxidants, leading to a disruption of redox signaling and control and/or molecular damage.

Oxidative Stress Definitions

The imbalance between oxidants and antioxidants in favor of the oxidants, potentially leading to damage, forms the core of the definition of oxidative stress. In the biological-biomedical field oxidative stress denotes a disturbance in the pro-oxidant–antioxidant balance in favor of the former. Much experimental evidence now supports the thesis that the favorable aspects of aerobic life are also linked to potentially dangerous oxygen-linked oxidation processes. These are thought to form the basis of a number of physiological and pathophysiological phenomena and to participate in processes as diverse as inflammation, atherosclerosis, aging, carcinogenesis, drug action and drug toxicity, and defense against protozoa. Research on oxidative stress has intensified in recent years, as documented by the close to 42,000 entries at the ISI Web of Knowledge at the beginning of 2005, mainly about the role of oxidants in cellular redox signaling.

In view of the proliferation of publications using the term, a few cautionary words might be appropriate. How can oxidative stress be defined operationally? In the normal healthy state, oxidative challenge occurs in many cell types, but this alone does not constitute an oxidative stress. Likewise, a simple loss of antioxidants resulting from a limited nutritional supply is not sufficient. However, when there is an increased formation of pro-oxidants such as hydrogen peroxide, which is accompanied by a loss of glutathione caused by the formation of glutathione disulfide (GSSG), we approach a definition. Even a severe loss of antioxidants, however, may still mean that there is no resulting damage. Novel roles of oxidants and antioxidants as part of signaling cascades and as mediators of adaptive responses have been unraveled, so it is possible that the time course and even the final outcome of a variety of diseases can be critically modulated by strengthening the antioxidant side of the balance.

This concept can be further refined by a better understanding of the normal biological uses of the reactive species that cause oxidative stress. The biological functions of reactive species that cause oxidative stress include fundamental processes such as the signaling of cell division and controlled cell elimination in the renewal of tissues. Furthermore, the balance of oxidants and antioxidants cannot be defined by a single entity. This has led to the view that multiple antioxidants are needed to protect cells and organs (the antioxidant network) and that different antioxidants are part of a protection system in specific disease processes. Following this line of reasoning, a refinement in the definition of oxidative stress has been suggested: "a disruption of redox signaling and control". Such a definition focuses attention on abnormal routes of electron flow, which both alter normal physiological signaling and damage the macromolecular machinery.

As these definitions indicate, oxidative stress is a concept that is becoming clearer with further scientific development, each ramification and extension providing valuable further input. Redox chemistry involves reduction–oxidation reactions, so obviously tipping the balance in favor of the reductants may also cause an imbalance (reductive stress). A landmark in the development of oxidative stress is the 1954 work of Gerschman, who viewed oxygen toxicity along with x-ray irradiation. A summary of the accumulated knowledge of the field appears in a 1979 review of hydroperoxide metabolism in mammalian organs by Chance et al. In a 1970 study of glutathione in red blood cells, the effect of glutathione reductase deficiency on the stimulation of a hexose monophosphate shunt under oxidative stress was examined by Beutler and colleagues; this may have been the first mention of the term, at least in the title of a publication that is readily retrievable.

Related Terms

Dietary Oxidative Stress

The term dietary oxidative stress was coined by Levander in 1995. The relationship between oxidative stress and viral infections has been scrutinized in regard to this concept. This area has received much attention from food science and in functional food considerations. In a definition of dietary antioxidants published by the Food and Nutrition Board of the Institute of Medicine, National Academy of Sciences, in 2000, the term was defined as "as a substance in foods that significantly decreases the adverse effects of reactive species, such as reactive oxygen species, reactive nitrogen species, on normal physiological function in humans." The ability of dietary anti- and pro-oxidants to affect oxidative stress has been widely discussed; Ames et al. have emphasized the presence of carcinogens and anticarcinogens in food.

Nutritional, or dietary, oxidative stress denotes a disturbance of the redox state resulting from excess oxidative load or from an inadequate nutrient supply that favors pro-oxidant reactions. The low intake of dietary antioxidants, including vitamins E and C, carotenoids, polyphenols, and other micronutrients (e.g., selenium), or the impaired availability of these antioxidants weakens the antioxidant network.

Postprandial oxidative stress, subform of nutritional oxidative stress, ensues from sustained postprandial hyperlipidemia and/or hyperglycemia and is associated with a higher risk for atherosclerosis, diabetes, and obesity. In Western societies, a significant part of the day is spent in the postprandial state. Unsaturated fatty acids incorporated into low-density lipoproteins (LDLs) and oxidized LDLs are an atherogenic factor. Lipid hydroperoxides present in the diet are absorbed and contribute to the pro-oxidant load. In hyperlipidemic and hyperglycemic subjects, endothelium-dependent vasodilation is impaired in the postprandial state, making postprandial oxidative stress an important factor modulating cardiovascular risk. Postprandial oxidative stress is attenuated when dietary antioxidants are supplied together with a meal rich in oxidized or oxidizable lipids. The ingestion of dietary polyphenols (e.g., from wine, cocoa, or tea) improves endothelial dysfunction and lowers the susceptibility of LDL lipids to oxidation. Polyphenols affect endothelial function not solely as antioxidants but also as modulatory signaling molecules.

Physiological Oxidative Stress

The fact that oxygen and its metabolites are not evenly distributed in organs and in subcellular organelles has led to the consideration that some sites are under physiological oxidative stress. The mitochondria are considered to be a site of physiological oxidative stress; in fact, it is possible that basically all cells, in the quasi-normal situation, experience this form of normal challenge. In view of the growing interest in adaptive responses, it is of interest to identify the need for and the extent of the physiological exposure to oxidative stress. The basal level of macrophage and neutrophil activation, for example, may fall into this category. Likewise, the physiological activity of glutamate (NMDA) receptors could be included here. Thus, in specific locations in the physiological system there is a routine deployment of oxidants for normal functioning. A difficulty arises, of course, in defining the level at which the threshold for the pathological oxidative stress is transcended. It may be that the accumulation of damage products, as part of the general definition of oxidative stress, could be a useful parameter in this regard. The adaptation to oxidative stress, initially studied predominantly with bacteria (e.g., in the oxyR-response) has also been demonstrated in eukaryotes and may have interesting further functions in the mammalian organism. The field of cell and tissues responses in terms of gene expression in mammalian systems is developing rapidly. The term exercise-induced oxidative stress may also fall under the present heading.

Spontaneous mutations and aging can be considered to be the result of the exposure of the organism to oxidative stress, be it physiological or pathophysiological. The loss of the *oxyR* regulon, which controls the expression of a battery of oxidant-defense genes, leads to an enormous increase in the number of spontaneous mutations. The relationship of oxidative stress to the aging process has been studied in many ways.

Photooxidative Stress

The generation of oxidants by electronic excitation through impinging light represents a major part of photobiology. The photoexcitation of endogenous or exogenous sensitizer molecules, leading to the formation of reactive oxygen species, notably singlet molecular oxygen, can be the cause of molecular and structural damage. The wavelength ranges of biological importance are ultraviolet B (UVB) and ultraviolet A (UVA), but visible light and even infrared (IR) are known to generate photobiological responses. In human health, exposed tissues such as skin and the eye are the most studied. There is nutritional protection against this damage from sunlight in dietary micronutrients such as carotenoids and polyphenols (flavonoids). Photodynamic therapy (PDT) uses the biological effects of photooxidative stress.

Radiation-Induced Oxidative Stress

The chemical basis of radiation biology involves the generation of free radicals. This is a large research field, ranging from chemistry to the treatment of cancer.

Oxidant Stress

Oxidant stress has been used much less frequently. This term (although it sounds slightly ambiguous) may be considered to be more or less synonymous with oxidative stress.

Pro-oxidant Stress

Pro-oxidant stress was used by Levander in his 1979 work on metal toxicity, again using red blood cells as a model, in describing the effects of lead. Cerutti used the term pro-oxidant states in his analysis of the role of oxidants in tumor promotion.

Oxidative Stress Status

The problem of assessing the biological status of intact systems has been tantalizing. Noninvasive methods are preferable, and stable end points are good candidates for useful applications. Our early work was on using GSSG efflux as an indicator. Many efforts were directed toward this (e.g., the use of 8-oxodG or certain isoprostanes in urinary samples). The study of biomarkers of oxidative stress is being actively pursued.

Approaches are evolving to evaluate oxidative stress and its association with the risk of disease in humans. Most methods are currently available only on an experimental basis. A key approach involves genetic analyses of variations in key antioxidant systems that have been associated with disease risk. Clinical measures of the status of the nutritionally required antioxidants vitamins C and E are available, but there is no consensus on the spectrum of additional antioxidants that should be measured for an effective evaluation of oxidative stress-related risk. The glutathione redox state and the determination of reactive oxygen metabolites provide more dynamic measures of the ongoing balance of pro-oxidant and antioxidant processes. In the blood plasma, the amino acid cysteine and its disulfide cystine are much more abundant than the glutathione system and provide an additional indicator of the balance of pro-oxidant and antioxidant systems. The cysteine/cystine balance in human plasma has been correlated with oxidative stress in aging, cigarette smoking, and cardiovascular disease. Although no consensus exists on ways to measure the damage from oxidative stress, measurement of isoprostanes, the free radical products of lipid oxidation found in the urine, by mass spectrometry is becoming a standardized and more generally accepted procedure. For the measurement of damage from reactive nitrogen species, the mass spectrometry measurement of nitrotyrosine in plasma proteins is being used as well.

Reductive Stress

In analogy to hyperoxia's being linked to oxidative stress, hypoxia or anoxia may be linked to reductive stress. The term reductive stress was introduced in 1987 by Wendel to denote the metabolic initiation of iron redox cycling following the overproduction of nicotinamide adenine dinucleotide (NADH). From a toxicological point of view, therefore, nicotinamide adenine dinucleotide phosphate (NADPH), or NADH, can be considered to be a detoxicant and a toxicant.

The effects caused by respiratory inhibition (chemical hypoxia), which favors the formation of toxic oxygen species, have been called reductive. The role of vascular endothelial growth factor in mediating the vascular dysfunction induced by hypoxia-like cytosolic metabolic imbalances, such as reductive stress and increased superoxide and nitric oxide production, has been studied. A gene battery responsive to reductive stress has been analyzed.

Nitrosative Stress

While working with nitrosothiols, Hausladen et al. observed a transcriptional activation of the *oxyR* regulon, which they called nitrosative stress. This area, which is of particular interest to vascular physiology and pathophysiology, has been treated in some detail.

Concluding Remarks

It is obvious that the term oxidative stress evokes different associations to researchers coming from different fields of interest. Thus, sometimes it has been overused, and caution is advised when the molecular species are not specified in the transition under study. The various concepts presented here may, in fact, not be at variance with one another; rather, the various terms emphasize a common phenomenon. Likewise, there are many parallelisms between the reactions distinguished as nitrosative versus oxidative. This shows that protection can be afforded against oxidative stress and nitrosative stress simultaneously, underlining the close relationship between the two.

Further Reading

Ames, B. N., Shigenaga, M. K. and Hagen, T. M. (1993). Oxidants, antioxidants, and the degenerative diseases of aging. *Proceedings of the National Academy of Sciences USA* **90**, 7915–7922.

Cerutti, P. A. (1985). Prooxidant states and tumor promotion. *Science* **227**, 375–381.

Chance, B., Sies, H. and Boveris, A. (1979). Hydroperoxide metabolism in mammalian organs. *Physiological Reviews* **59**, 527–605.

Cutler, R. G. and Rodriguez, H. (eds.) (2003). *Oxidative stress and aging*. Singapore: World Scientific Publishing.

Dalle-Donne, I., Scaloni, A., Giustarini, D., et al. (2005). Proteins as biomarkers of oxidative/nitrosative stress in diseases: the contribution of redox proteomics. *Mass Spectrometry Reviews* **24**, 55–99.

Food and Nutrition Board, Institute of Medicine (2000). Dietary reference intakes for vitamin C, vitamin E, selenium, and carotenoids, p. 17. National Academy Press: Washington.

Griffiths, H. R., Moller, L., Bartosz, G., et al. (2002). Biomarkers. *Molecular Aspects of Medicine* **23**, 101–208.

Halliwell, B. (1996). Oxidative stress, nutrition and health: experimental strategies for optimization of nutritional antioxidant intake in humans. *Free Radical Research* **25**, 57–74.

Hausladen, A., Privalle, C. T., Keng, T., et al. (1996). Nitrosative stress: activation of the transcription factor OxyR. *Cell* **86**, 719–729.

Hermez-Lima, M. (2004). Oxidative stress and medical sciences. In: Story, K. (ed.) *Functional metabolism: regulation and adaptation*, pp. 369–382. New York: John Wiley.

Jones, D. P., Mody, V. C. Jr., Carlson, J. L., et al. (2002). Redox analysis of human plasma allows separation of prooxidant events of aging from decline in antioxidant defenses. *Free Radical Biology & Medicine* **33**, 1290–1300.

Levander, O. A., Fontela, R., Morris, V. C., et al. (1995). Protection against murine cerebral malaria by dietary-induced oxidative stress. *Journal of Parasitology* **81**, 99–103.

Paniker, N. V., Srivastava, S. K. and Beutler, E. (1970). Glutathione metabolism of the red cells: effect of glutathione reductase deficiency on the stimulation of hexose monophosphate shunt under oxidative stress. *Biochimica et Biophysica Acta,* **215,** 456–460.

Pryor, W. A. and Godber, S. S. (1991). Noninvasive measures of oxidative stress status in humans. *Free Radical Biology & Medicine* **10**, 177–184.

Schwarz, K. B. (1996). Oxidative stress during viral infection: a review. *Free Radical Biology & Medicine* **21**, 641–649.

Sies, H. (1985). Oxidative stress: introductory remarks. In: Sies, H. (ed.) *Oxidative stress*, pp. 1–8. London: Academic Press.

Sies, H. (1986). Biochemistry of oxidative stress. *Angewandte Chemie International Edition in English* **25**, 1058–1071.

Sies, H. (ed.) (1997). *Antioxidants in disease mechanisms and therapy.* San Diego: Academic Press.

Sies, H. (2000). What is oxidative stress? In: Keaney, J. F. (ed.) *Oxidative stress and vascular diseases*, pp. 1–8. Boston: Kluwer Academic.

Sies, H. and Stahl, W. (2004). Nutritional protection against skin damage from sunlight. *Annual Review of Nutrition* **24**, 173–200.

Sies, H., Stahl, W. and Sevanian, A. (2005). Nutritional, dietary and postprandial oxidative stress. *Journal of Nutrition* **135**, 969–972.

Sohal, R. S. and Weindruch, R. (1996). Oxidative stress, caloric restriction, and aging. *Science* **273**, 59–63.

Storz, G., Christman, M. F., Sies, H., et al. (1987). Spontaneous mutagenesis and oxidative damage to DNA in *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences USA* **84**, 8917–8921.

von Sonntag, C. (1987). *The chemical basis of radiation biology.* London: Taylor & Francis.

Wendel, A. (1987). Measurement of *in vivo* lipid peroxidation and toxicological significance. *Free Radical Biology & Medicine* **3**, 355–358.

Wink, D. A., Cook, J. A., Kim, S. Y., et al. (1997). Superoxide modulates the oxidation and nitrosation of thiols by nitric oxide-derived reactive intermediates: chemical aspects involved in the balance between oxidative and nitrosative stress. *Journal of Biological Chemistry* **272**, 11147–11151.

Oxidative Stress and Acidosis, Molecular Responses to

C Mobbs
Mt. Sinai School of Medicine, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C Mobbs, volume 3, pp 105–108, © 2000, Elsevier Inc.

Oxidative Stress and Acidosis
Molecular Responses to Oxidative Stress and Acidosis
Mechanisms Mediating Molecular Responses to Oxidative Stress and Acidosis
Summary

Glossary

Acidosis	An increase in the acidity, and thus a decrease in pH.
Cysteine	An amino acid (subunit of protein) particularly sensitive to oxidation.
Cytochrome B	An essential component of mitochondrial metabolism, which functions in electron transport chain complex III.
Cytochrome oxidase	An essential component of mitochondrial metabolism, which functions to facilitate electron transport.
Free radical	A molecular species such as O_2^-, which by virtue of a free electron is energetically favored to reactively oxidize many compounds and cellular constituents.
Heme oxidase 1	An enzyme that protects cells from oxidative damage.
Metallothionein	A protein that protects cells from oxidative damage.
Mitochondrial rRNAs	Ribosomal RNAs used to translate proteins in the mitochondria.
Oxidation	A chemical reaction in which a compound or molecular constituent loses electrons.
Oxidative state	A reflection of the likelihood that atoms in the environment will gain or lose net electrons or hydrogen atoms; in a more highly oxidative state sulfur atoms in the side chains of the amino acid cysteine will tend to lose net electrons and hydrogen atoms, thus enhancing the production of disulfide bridges.
Oxidative stress	Potentially toxic increase in oxidative state of the cell, leading to characteristic molecular responses.
OxyR protein	A protein that regulates gene expression which, when activated by oxidative stress, stimulates the expression of antioxidant genes.
Reduction	A chemical reaction in which a compound or molecular constituent gains electrons.
Superoxide dismutase	An enzyme that protects against oxidative stress by inactivating the highly oxidant free radical O_2^-.

Oxidative Stress and Acidosis

The internal milieu of the cell is highly regulated, and any displacement from this milieu can have fatal consequences. Therefore molecular mechanisms have evolved to counter environmental influences which threaten to disrupt this milieu. These molecular mechanisms entail responses which are common (stereotypic) in a wide range of cells and in response to a wide variety potentially toxic influences. Thus, these molecular responses may usefully be considered to be molecular stress responses. However, these general stress responses are also accompanied by more specific changes which depend on the cell type and specific stressor. Perhaps the best-studied of these molecular stress responses are the heat-shock proteins, which were first observed to be induced by a potentially lethal increase in temperature. The heat-shock proteins are discussed in detail elsewhere in this encyclopedia. As discussed there, the heat-shock proteins are not only induced by elevated temperature, but also by a wide variety other stressors, among which are changes in the oxidative state and the pH of the cell. However, oxidative stress and acidosis also induce other molecular responses which are not prominent responses to heat shock. These other responses are the focus of this present article.

Oxidative state can best be thought of as the relative attraction of electrons to atoms or compounds. Oxygen is one of the strongest attractors of electrons, so it is relatively oxidizing comparing to almost any other atom. In comparison, hydrogen has relatively little attraction for electrons. Thus when hydrogen and oxygen combine to form water, the hydrogen is oxidized by (loses electrons to) oxygen, and, conversely, the oxygen is reduced by (gains electrons from) hydrogen. In a more general sense, it may be said that the atmosphere, consisting of a relatively high level of oxygen due to ongoing photosynthesis, is characterized by a relatively high oxidative state compared, for example, to the internal state of the cell, in which free oxygen is at much lower concentrations. This disparity partly reflects the fact that life originally evolved in

atmospheric conditions which were much less oxidative than exists today. The physiological significance of this disparity is that the proper function of most proteins requires an oxidative state lower than is found in outside the cell. Even within the cell, some compartments, such as the endoplasmic reticulum, require even more highly reduced conditions than other parts of the cell. However, most cells require oxygen to produce energy, and in the course of producing energy molecular species called free radicals are produced which potentially greatly enhance the local oxidative state inside the cell. Furthermore, other cellular functions, such as certain immunological functions, may entail enhanced production of free radicals. When free radicals or oxidative chemicals increase the oxidative state inside the cell, the result can be the oxidation of proteins, especially at disulfide bonds that are highly sensitive to oxidative status. Because such an increase in oxidative state may be toxic and induces a common set of molecular responses, this change is called oxidative stress. In principle cellular function could also become impaired if oxidative state becomes too low (that is, the internal milieu could be too reduced), but this circumstance rarely occurs in most cells and is not further considered in this article.

In addition to the highly regulated oxidative state, the acidity, or pH, inside the cell is also highly regulated. Acidity, which reflects the concentration of free hydrogen ions, is critically important because the ionic status of most amino acid side chains in proteins is dependent on pH. Therefore, agents that alter internal pH can, like agents that alter oxidative status, also be toxic, and thus not surprisingly cellular responses to acidosis have also evolved. Much less is known about these responses, but molecular responses to acidosis and oxidative stress entail induction of some common species.

Molecular Responses to Oxidative Stress and Acidosis

Molecular responses to oxidative stress entail some common responses and some responses that depend on the cell type and nature of the oxidative stressor. For example, as with many potentially toxic insults, oxidative stress may generally entail induction of heat-shock proteins. However, other responses are different from those induced by heat shock. The range of molecular responses to oxidative stress is indicated by comparing three papers that used the differential display technique to assess molecular responses of three different cell types to either hydrogen peroxide or diethylmaleate. Hydrogen peroxide directly increases oxidative state since it is a highly oxidative molecule, whereas diethylmaleate increases the oxidative state by depleting the reduced form of glutathione. In a Chinese hamster fibroblast cell line, hydrogen peroxide was observed to cause a decrease in a several transcripts derived from mitochondria (including both 16S and 12S mitochondrial rRNAs, nicotinamide adenine dinucleotide (NADH) dehydrogenase subunit 6; ATPase subunit 6; and cytochrome oxidase subunits I and III). In contrast, most nucleus-derived ribonucleic acid (RNA) species were either not influenced or increased, including messenger RNAs (mRNAs) for several growth-inhibiting proteins. In a human lens epithelial cell line, hydrogen peroxide was observed to influence transcripts derived from both mitochondrial and nuclear genomes. The mitochondrial transcripts (including NADH dehydrogenase subunit 4, cytochrome-b, cytochrome oxidase III, and 16S RNA) were decreased. In contrast, all but one of the nucleus-derived transcripts that were regulated (cytokine-inducible nuclear protein, alternative splicing factor 2, and beta-hydroxyisobutyryl-coenzyme A hydrolase) were increased. Among the nucleus-derived transcripts, only glutamine cyclotransferase was decreased. In the human-derived Hela cell line, diethylmaleate was observed to decrease the levels of RNA species derived from mitochondria, whereas nuclear-derived species, including the ribosomal protein 4, c-fos, and previously uncharacterized mRNAs, were induced. Thus the differential display technique, which in effects surveys across all RNA species, clearly indicates that a common response to oxidative stress is a decrease in mitochondrial RNA species and an induction of a small number of nuclear-derived mRNAs. Since oxidative stress would normally be produced by mitochondrial oxidative metabolism, the decrease in mitochondrial species can be understood as a homeostatic response to decrease the production of oxidative stress. This common oxidative stress response, however, is not observed with heat shock. In contrast, other proteins, including antioxidant proteins such as glucose-6-phosphate dehydrogenase (which catalyzes the rate-limiting step in the production of nicotinamide adenine dinucleotide phosphate hydrogen (NADPH), the source of reducing potential for the antioxidant glutathione), catalase, superoxide dismutase, heme oxidase 1, and metallothionein, are induced by many forms of oxidative stress in many cell types. While these gene products may be induced by thermal or other stressors under some circumstances, their induction by oxidative stress is much more robust. Interestingly, other than heat-shock proteins, the most robust reported response to acidosis is also metallothionein. Both metallothionein and heme oxidase 1 probably serve as antioxidants

by sequestering potentially oxidizing metals, but this remains to be established.

The molecular responses to oxidative stress depend on the severity of the stress, in relation to whether the stress is lethal, produces programmed cell death (apoptosis), or is relatively mild and not only allows the cell to survive but become relatively resistant to subsequent cell death. DNA microarray studies suggest that relatively mild oxidative stress produce molecular responses that could lead to apoptosis, but these potentially lethal responses are inhibited, interestingly, by a concomitant reduction in glycolysis. This reduction in glycolysis actually appears to produce a protective effect, in part by inhibiting apoptosis, but through other mechanisms as well. However, if the oxidative stress becomes severe enough, the apoptotic effects overwhelm the blockade produced by glycolysis inhibition. Even more severe oxidative stress produces cell death directly, without apoptosis (necrosis).

Mechanisms Mediating Molecular Responses to Oxidative Stress and Acidosis

In eukaryotic systems, oxidative stress appears to lead to degradation of mitochondrial RNA species through unknown mechanisms. Oxidative stress increases nuclear-derived mRNAs through several mechanisms, including increased transcription and, as for heme oxidase 1, increased mRNA stability. The molecular details of these effects of oxidative stress in eukaryotes is not well-understood. However, oxidative stress responses in prokaryotes have been examined in some detail.

In bacteria, hydrogen peroxide induces about 30–40 proteins, whereas agents which lead to free production induce an approximately 40 additional proteins. Hydrogen peroxide induces gene expression in part by activating the OxyR protein, which activates transcription of about 10 antioxidant proteins including catalase (the enzyme which converts hydrogen peroxide to water), and glutathione reductase. OxyR is normally inactive in the relatively reduced normal cellular milieu, but hydrogen peroxide leads to an oxidation of two cysteine residues on OxyR which activates the protein. Free radical production induces gene expression in part through the oxidation of SoxS, a transcriptional activator which stimulates the transcription of several antioxidant genes including superoxide dismutase and glucose-6-phosphate dehydrogenase. The oxidation of SoxS takes place at an iron–sulfur center, so requires a higher energy form of oxidation than is produced by hydrogen peroxide alone. Thus, molecular responses to different forms of oxidative stress can be appropriately generated through effector molecules with sensitivities to different forms of oxidation. Recent studies suggest that a similar process occurs in eukaryotes, entailing the activation of the transcription factor Nrf2 via the mitogen-activated protein (MAP) kinase signaling pathway.

Summary

Life initially evolved under conditions of much lower oxygen than is the case at present and most cellular processes therefore require a lower state of oxidation than is found in air. Therefore specific and highly regulated processes have evolved to maintain this lower state of oxygen even in the face of metabolic insults. Several levels of antioxidant defenses exist, to accommodate precise levels of oxidative challenges. Some of these antioxidative proteins, including metallothioneins, are also induced to protect against acidosis.

Further Reading

Ammendola, R., Fiore, F., Esposito, F., et al. (1995). Differentially expressed mRNAs as a consequence of oxidative stress in intact cells. *FEBS Letters* **371**, 209–213.

Colussi, C., Albertini, M. C., Coppola, S., et al. (2000). H_2O_2-induced block of glycolysis as an active ADP-ribosylation reaction protecting cells from apoptosis. *Faseb Journal* **14**, 2266–2276.

Carper, D. A., Sun, J. K., Iwata, T., et al. (1999). Oxidative stress induces differential gene expression in a human lens epithelial cell line. *Investigative Ophthalmology & Visual Science* **40**, 400–406.

Choi, A. M. and Alam, J. (1996). Heme oxygenase-1: function, regulation, and implication of a novel stress-inducible protein in oxidant-induced lung injury. *American Journal of Respiratory Cell and Molecular Biology* **15**, 9–19.

Conklin, D. R., Tan, K. H. and Aschner, M. (1998). Dimethyl sulfoxide, but not acidosis-induced metallothionein mRNA expression in neonatal rat primary astrocyte cultures is inhibited by the bioflavonoid, quercetin. *Brain Research* **794**, 304–308.

Crawford, D. R., Wang, Y., Schools, G. P., et al. (1997). Down-regulation of mammalian mitochondrial RNAs during oxidative stress. *Free Radical Biology & Medicine* **22**, 551–559.

Demple, B. (1999). Radical ideas: genetic responses to oxidative stress. *Clinical and Experimental Pharmacology and Physiology* **26**, 64–68.

Narasimhan, P., Swanson, R. A., Sagar, S. M., et al. (1996). Astrocyte survival and HSP70 heat shock protein induction following heat shock and acidosis. *Glia* **17**, 147–159.

Nebert, D. W., Petersen, D. D. and Fornace, A. J., Jr. (1990). Cellular responses to oxidative stress: the [Ah] gene

battery as a paradigm. *Environmental Health Perspectives* **88**, 13–25.
Nishimura, R. N., Dwyer, B. E., Cole, R., et al. (1989). Induction of the major inducible 68-kDa heat-shock protein after rapid changes of extracellular pH in cultured rat astrocytes. *Experimental Cell Research* **180**, 276–280.
Pahl, H. L. and Baeuerle, P. A. (1994). Oxygen and the control of gene expression. *Bioessays* **16**, 497–502.
Piette, J., Piret, B., Bonizzi, G., et al. (1997). Multiple redox regulation in NF-kappaB transcription factor activation. *Biological Chemistry* **378**, 1237–1245.
Ryter, S. W. and Choi, A. M. (2005). Heme oxygenase-1: redox regulation of a stress protein in lung and cell culture models. *Antioxidants & Redox Signaling* **7**, 80–91.
Schreck, R., Albermann, K. and Baeuerle, P. A. (1992). Nuclear factor kappa B: an oxidative stress-responsive transcription factor of eukaryotic cells (a review). *Free Radical Research Communications* **17**, 221–237.
Tawe, W. N., Eschbach, M. L., Walter, R. D., et al. (1998). Identification of stress-responsive genes in *Caenorhabditis elegans* using RT-PCR differential display. *Nucleic Acids Research* **26**, 1621–1627.
Zhang, Y., Fong, C. C., Wong, M. S., et al. (2005). Molecular mechanisms of survival and apoptosis in RAW 264.7 macrophages under oxidative stress. *Apoptosis* **10**, 545–556.

Oxidative Stress and Aging

C Mobbs
Mt. Sinai School of Medicine, New York, USA

© 2007 Elsevier Inc. All rights reserved.

Metabolism Drives Aging: The Free Radical Hypothesis
Oxidative Stress Increases with Age
Functional Evidence That Oxidative Stress Limits Life Span

Glossary

Free radical	A molecular species such as O_2^-, which by virtue of a free electron is energetically favored to reactively oxidize many compounds and cellular constituents.
Mitochondria	The powerhouse of the cell, where ATP, the main source of cellular energy, is produced.
Mitochondrial electron transfer chain	A specialized process in the mitochondria whereby electrons are passed down a chain of five complexes, allowing for the production of ATP.
Oxidation	A chemical reaction in which a compound or molecular constituent loses electrons.
Oxidative state	A reflection of the likelihood that atoms in the environment will gain or lose net electrons or hydrogen atoms; in a more highly oxidative state, sulfur atoms in the side chains of the amino acid cysteine tend to lose net electrons and hydrogen atoms, thus enhancing the production of disulfide bridges.
Oxidative stress	The potentially toxic increase in oxidative state of the cell, leading to characteristic molecular responses.
Reduction	A chemical reaction in which a compound or molecular constituent gains electrons.

Metabolism Drives Aging: The Free Radical Hypothesis

The causes of aging have been the subject of speculation since before the dawn of recorded history, and in the era of science such speculations have become even more numerous. At present, to the extent that a unitary mechanism of sorts is seriously entertained by gerontologists, the hypothesis that aging is caused by oxidative damage is the most widely accepted. This hypothesis had its origins in the observation that life span seems inversely correlated with metabolic rate, which gave rise, early in the twentieth century, to the rate of living theory. This hypothesis, inspired by the fantastically successful conservation laws of classical physics, hypothesized that the total amount of metabolic activity an individual could sustain over its life span is a constant. This hypothesis accorded well with the basic observation with inanimate objects that use leads to deterioration and with the intuition that, in human life at least, those who live fast, die young; it also accorded well with a more objective analysis that, for example, heart rate is much higher in short-lived mammals than in long-lived mammals. Indeed, this observation gave rise to the hypothesis that the lifetime number of heartbeats, thought to be a proxy of metabolic rate, is conserved. The hypothesis was strengthened by the incontrovertible demonstration that reducing the temperature in

poikolotherms such as flies and nematodes increased life span. Because reducing the temperature *in vitro* reduces the rate of metabolic and chemical reactions, these observations were highly suggestive that metabolic processes drive aging and limit life span.

Nevertheless, precisely which aspects of metabolism limits life span has continued to be elusive. Recent studies have suggested that total metabolic activity is in fact not conserved; for example, caloric restriction can increase life span without producing a concomitant decrease in tissue-specific metabolic rate. In 1956, by analogy with radiation-induced cellular damage, Dennis Harman suggested that the metabolic process that limits life span is the production of free radicals whose main damaging effect is mediated through the oxidation of macromolecules such as protein and DNA. This free-radical hypothesis of aging emphasized the role of oxidative stress, although in principle oxidative stress could occur through other mechanisms in addition to free radicals. This hypothesis has continued to stimulate much research and continues to be among the most attractive mechanisms to mediate the aging process.

Oxidative stress, defined as evidence of a relatively oxidized state of the cell and its constituents, has been implicated in the aging process by several lines of evidence. Originally, this hypothesis was suggested by the observation that radiation, which can certainly produce oxidative stress, can in some ways mimic the aging process, especially in reducing life span and increasing the incidence of tumor burden. On the other hand, it is not certain that oxidative damage necessarily mediates the pathological effects of radiation, and it has been similarly difficult to prove that oxidative stress mediates the pathologies of aging. Nevertheless, evidence implicating oxidative stress as mediating the pathological effects of aging has continued to grow over the last 50 years.

Oxidative Stress Increases with Age

Perhaps the most obvious evidence is that oxidative stress incontrovertibly increases with age. For example total protein oxidation, as indicated by protein carbonyl levels and nitrotyrosine immunoreactivity, increases by approximately 50% in the brain from maturity to old age. Conversely, total protein reduction, as indicated by sulfhydryl groups, decreased with age. For some proteins, such as glyceraldehyde phosphate dehydrogenase, the age-related oxidation has been located at specific cysteine residues, and the oxidation has been shown to account for the robust decrease in specific activity characteristic of many enzymes during aging. Furthermore, protein oxidation in specific brain regions of aged mice has been shown to correlate with impairments in functions regulated by those specific brain regions. Thus, for example, the oxidation of hippocampal proteins correlates with impairments in memory function, whereas the oxidation of cerebellar proteins correlates with impairments in tasks requiring fine motor coordination. Oxidative damage to DNA similarly increases with age, especially in mitochondrial DNA. In mitochondria, DNA oxidative damage is mainly manifested by characteristic deletions that can entail the loss of several thousand bases and whose occurrences increases robustly with age in many tissues. In nuclei, DNA oxidation is mainly manifest as the modification of specific bases, especially the conversion of guanine to 8-oxoguanine, and specific oxidized bases in the nucleus, which are most likely due to oxidative damage. Finally, the oxidation of lipids, mainly manifest in the form of 4-hydroxynonenal (HNE) and malondialdehyde, also increases with age in several tissues.

The increased oxidative damage that occurs during aging appears to be due primarily to the increased production of reactive oxygen species during aging, although reduced turnover of proteins and lipids may also play a role. Most reactive oxygen species are produced by the mitochondria, especially in electron transport chain complexes I and III. It is not clear why the production of reactive oxygen species increases with age, but in some tissues at least this may be due to a shift toward the greater use of complex I and away from the use of complex II.

Functional Evidence That Oxidative Stress Limits Life Span

Four main lines of evidence suggest that the increase in oxidative damage that occurs during aging plays an important role in determining life span. First, age increases the lethality of oxidative stress, for example in *Caenorhabditis elegans*. Second, dietary restriction, which increases life span, also increases resistance to oxidative stress. Third, genetic modifications that increase life span (e.g., mutations in the insulin-like receptor) also increase resistance to oxidative stress. Fourth, genomewide unbiased screening revealed that the most prominent class of genes whose ablation increases life span are genes coding for mitochondrial electron transport chain complexes I, III, IV, and V. As previously indicated, these complexes, especially I and III, are the major sources of reactive oxygen species in most cells. Notably missing from these screens were genes coding for proteins in mitochondrial electron transport chain complex II. Indeed, it had already been shown by classic genetic screens that mutations in complex II reduced life span rather than

increasing it. The significance of these observations is that reactive oxygen species are produced at much greater levels in complex I than in complex II. Thus, mutations that reduce complex II activity presumably entail elevated compensatory activity of complex I, increasing the production of reactive oxygen species and reducing life span, whereas genetic manipulations that inhibit complex I have the opposite effect. Precisely such a mechanism seems to mediate the age-related pathologies (and reduced life span) caused by Huntington's disease.

A final line of evidence supporting a role for oxidative stress in life span involve studies that genetically manipulate genes coding for the classic anti-oxidative defense enzymes superoxide dismutase and catalase. Although early studies in flies suggested that transgenic overexpression of these genes might increase life span, later studies failed to replicate this effect. Furthermore, in at least one study genetic reduction in superoxide dismutase activity had no effect on life span, even though this manipulation increased oxidative stress as well as tumor burden. Therefore, the effect of manipulating the activity of classic anti-oxidant defenses on life span must be considered an open question. On the other hand, these classic anti-oxidant enzymes are not the sole determinant of oxidative stress, so the best interpretation of the current evidence is that oxidative stress, but not necessarily the classic anti-oxidant enzymes, is a major determinant of age-related pathologies and mortality.

Further Reading

Butterfield, D. A., Drake, J., Pocernich, C., et al. (2001). Evidence of oxidative damage in Alzheimer's disease brain: central role for amyloid beta-peptide. *Trends in Molecular Medicine* **7**(12), 548–554.

Ishii, N., Goto, S. and Hartman, P. S. (2002). Protein oxidation during aging of the nematode *Caenorhabditis elegans*. *Free Radical Biology & Medicine* **33**(8), 1021–1025.

Lee, S. S., Lee, R. Y., Fraser, A. G., et al. (2003). A systematic RNAi screen identifies a critical role for mitochondria in *C. elegans* longevity. *Nature Genetics* **33**(1), 40–48.

Lin, Y. J., Seroude, L. and Benzer, S. (1998). Extended life-span and stress resistance in the Drosophila mutant methuselah. *Science* **282**(5390), 943–946.

Lopez-Torres, M., Gredilla, R., Sanz, A., et al. (2002). Influence of aging and long-term caloric restriction on oxygen radical generation and oxidative DNA damage in rat liver mitochondria. *Free Radical Biology & Medical* **32**(9), 882–889.

Nicolle, M. M., Gonzalez, J., Sugaya, K., et al. (2001). Signatures of hippocampal oxidative stress in aged spatial learning-impaired rodents. *Neuroscience* **107**(3), 415–431.

Perez-Campo, R., Lopez-Torres, M., Cadenas, S., et al. (1998). The rate of free radical production as a determinant of the rate of aging: evidence from the comparative approach. *Journal of Comparative Physiology B* **168**(3), 149–158.

Sohal, R. S., Mockett, R. J. and Orr, W. C. (2002). Mechanisms of aging: an appraisal of the oxidative stress hypothesis. *Free Radical Biology & Medicine* **33**(5), 575–586.

Wallace, D. C. (2005). A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. *Annual Review of Genetics* **39**, 359–407.

Yuh, K. C. and Gafni, A. (1987). Reversal of age-related effects in rat muscle phosphoglycerate kinase. *Proceedings of the National Academy of Sciences USA* **84**(21), 7458–7462.

Oxytocin

G Leng and N Sabatier
University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Leng, volume 3, pp 109–114, © 2000, Elsevier Inc.

Effects of Stress on the Milk-Ejection Reflex
Stress and Parturition
Stress-Induced Stimulation of Oxytocin Secretion
Oxytocin Release within the Brain
Conclusion

Glossary

Dendrites	Parts of a cell where information is received and processed; thick, specialized processes arising from the cell body (soma) of a nerve cell (neuron). Most inputs to a nerve cell from other nerve cells arrive at the dendrites. However, for many cells, the transfer of information at the dendrites is bidirectional; neurotransmitters and neurohormones secreted from the dendrites can affect the activity of neighboring cells and can also affect the excitability of the afferent nerve endings.

Natriuresis Sodium excretion in urine. To keep the plasma concentration of sodium stable, the intake of salt in the diet has to be balanced with appropriately regulated excretion. Natriuresis is regulated by a variety of neural and hormonal mechanisms.

Posterior pituitary gland The site at which oxytocin and vasopressin are secreted into the general circulation; also called the neurohypophysis. The gland contains only the axons and neurosecretory nerve terminals of magnocellular vasopressin and oxytocin cells; the cell bodies are within the hypothalamus.

Oxytocin, a peptide hormone, is synthesized by large (magnocellular) neurons in the supraoptic and paraventricular nuclei of the hypothalamus. The axons of these neurons project to the posterior pituitary gland, where oxytocin is stored in large vesicles within nerve endings and axon swellings. Electrical activity in the cell bodies is propagated down the axons in the form of action potentials, resulting in the secretion of oxytocin, which then enters the circulation. Thus, the concentration of oxytocin in the blood reflects the electrical activity of oxytocin neurons in the hypothalamus. There are also some peripheral sources of oxytocin, but these are likely to play local, autocrine roles and do not make a substantial contribution to concentrations in the circulation.

Effects of Stress on the Milk-Ejection Reflex

Oxytocin is essential for the transfer of milk from mother to infant. The sucking of a baby at a mother's breast activates sensory receptors in and around the nipple, leading to the activation of spinal nerves. These nerves project indirectly to the hypothalamus, leading to the activation of oxytocin neurons and secretion of oxytocin. The oxytocin is carried in the systemic circulation to the mammary gland, where it acts on myoepithelial cells, causing milk to be let down from these cells into the collecting duct of the mammary gland, from which the baby can draw the milk into his or her mouth. Without oxytocin secretion, there is no milk let-down and a suckling baby will go hungry.

This reflex – the milk-ejection reflex – is abrupt and intermittent. In response to suckling, oxytocin is secreted in pulses every few minutes. These pulses are the result of brief, intense, coordinated bursts of electrical activity in the oxytocin cells, separated by periods of relative electrical quiescence. Although the reflex is normally activated by suckling, it can also be conditioned – and the sight, sound, or smell of a baby can, in some circumstances, elicit milk let-down.

This reflex is very susceptible to disruption by stress. In humans, as in many mammals, a quiet, familiar, and safe environment and a calm state of mind are conducive to the reflex. Indeed, in rats the milk-ejection reflex appears to require that cortical activity be in a state of slow-wave sleep. However, although stress might thus be described as inhibiting the oxytocin-mediated milk-ejection reflex, it does not follow that stress inhibits oxytocin secretion. The pulsatile pattern of oxytocin secretion in response to suckling means that the mammary gland is exposed to very high circulating concentrations of oxytocin – but for only a few seconds at a time. The gland responds only to high concentrations of oxytocin, and continuous exposure to high concentrations soon leads to desensitization, so anything that disrupts the appropriate patterning of oxytocin secretion can block the milk-ejection reflex. The reflex requires a coordinated, abrupt activation of the oxytocin cells, and many stimuli that increase the basal activity of oxytocin cells also impair their ability to display coordinated, abrupt episodes of activation. Hence, perversely, many stimuli that lead to increased concentrations of oxytocin in the blood at the same time block the pulsatile pattern of secretion that is essential for the physiological efficacy of oxytocin at the mammary gland.

Stress and Parturition

Oxytocin plays an important role during parturition in most, if not all, mammals. In all mammals studied, the uterus becomes highly sensitive to oxytocin toward the end of pregnancy, as a result of the increased expression of oxytocin receptors in both the myometrium and endometrium. Oxytocin acts on the myometrium of the pregnant uterus to induce contractions directly and on the endometrium to stimulate the production of prostaglandins, which also induce uterine contractions. During delivery, the stimulation of the cervix initiates the Ferguson reflex, during which large amounts of oxytocin are secreted from the pituitary gland into the circulation and uterine contractions and abdominal compression also contribute to the activation of oxytocin secretion.

In the rat, oxytocin cells behave during parturition in much the same way as during suckling – they display intermittent intense bursts of activity leading to pulsatile secretion of hormone. Again, this pattern is important for the physiological efficacy of

oxytocin, although less so than in the case of the milk-ejection reflex; the uterus is sensitive to lower concentrations of oxytocin in the circulation and is less susceptible to desensitization.

Like many mammals, including the rabbit and the sow, the pregnant rat builds a nest in which to deliver; some other mammals, such as the cat, ewe, and bitch, do not build a nest but instead seek out a sheltered place. Mammals that are normally most active at night tend to give birth during the day, and those most active during the day tend to give birth at night; in women, the peak incidence for delivery is between 3:00 and 4:00 a.m. Thus, labor generally begins at a time when the animal is normally asleep or resting and in a place that offers relative seclusion and apparent safety. Disturbance – the sight or sound of an intruder or physical movement from the place where labor has begun to a new environment – can disrupt parturition markedly. If a mouse or a rat is taken from her nest once birth has started and is placed in an empty glass chamber or if a sow is moved from the farrowing nest that she has made to an empty pen, subsequent deliveries are delayed. In bitches, ferrets, rabbits, and ewes, even the noticed presence of a human observer can result in prolonged labor associated with the inhibition of uterine contractions. In women too, anxiety is associated with prolonged labor, and although this has not been the subject of controlled observations, uterine contractions often decrease temporarily when a woman in labor is moved from home to the hospital or from the labor room to the delivery room.

Thus, the progress of parturition in mammals is very susceptible to disruption by acute stress, and many authors have suggested that this disruption might be a consequence of a disruption of oxytocin secretion. The acute administration of oxytocin antagonists prolongs parturition in the rat, and the pharmacological inhibition of oxytocin secretion, particularly by opioid agonists, leads to a prolongation of parturition that can be reversed by the administration of physiological doses of oxytocin. Thus, if stress impairs the activation of oxytocin neurons in parturition, then this would be likely to impair the progress of parturition.

When a rat is moved in mid-parturition from her nesting cage to an empty glass bowl, the oxytocin concentrations in the blood fall, but whether this is a primary consequence of stress or is secondary to reduced uterine activity is not clear. In the rat and in the sow, after the administration of the opioid antagonist naloxone, parturition progresses normally despite environmental disturbance, suggesting that stress might activate endogenous opioid pathways that inhibit oxytocin secretion; oxytocin neurons are known to be very sensitive to inhibition by exogenous opioids. However, in the rat, naloxone administered during parturition potentiates oxytocin secretion even in the absence of overt stress, so it is possible that the normal progress of parturition in the naloxone-treated disturbed rat merely reflects the ability of hypersecretion of oxytocin to overcome the influence of mild stress.

Thus, mild stress can subvert the physiological actions of oxytocin, preventing milk let-down in response to suckling and prolonging parturition. In lactation, stress disrupts the pattern of oxytocin secretion, preventing the coordinated intermittent burst discharge of oxytocin cells that is necessary for pulsatile oxytocin secretion. During parturition, it is less clear how stress interferes with the progress of labor, but it seems plausible that a similar disruption of the patterning of oxytocin secretion leads to a less effective hormonal signal at the uterus. In the case of the milk-ejection reflex, the disruption of the appropriate patterning of secretion might be accompanied by (and might even be caused by) an increased but uncoordinated activation of the oxytocin cells, giving rise to higher average concentrations of oxytocin in the circulation.

Stress-Induced Stimulation of Oxytocin Secretion

In the rat, on which most studies have been performed, most acute stressors that have been tested markedly increase oxytocin secretion from the pituitary gland. Plasma concentrations of oxytocin are increased, for instance, following forcible restraint, noxious stimuli, cold exposure, forced swimming, aversive foot shocks, administration of nausea-inducing agents, and ether inhalation. At least some of the neural pathways that are involved are also involved in mediating the release of corticotropin releasing hormone (CRH) and hence the secretion of adrenocorticotropic hormone (ACTH). However, stressors described as emotional only, such as social defeat, do not induce oxytocin secretion. Oxytocin secretion also depends on the species; plasma concentrations of oxytocin do not change in humans exposed to stress or in horses exposed to a novel environment. Stress-induced oxytocin secretion in rats is partly mediated by A2 and A1 noradrenergic cells of the brain stem because the oxytocin response to fear-related stimuli or noxious stimuli, for instance, is impaired after central norepinephrine depletion by neurotoxin treatment or by central injection of an α_1-adrenergic antagonist. However, not all stressors are mediated by

noradrenergic pathways; for example, the disruption of noradrenergic pathways does not affect oxytocin secretion in rats exposed to a novel environment.

These observations raise several questions. Is the stress-induced secretion of oxytocin into the circulation of any physiological significance? Or is it an epiphenomenon, reflecting a disruption of the normal pattern of activity in oxytocin cells that is essential for reflex milk ejection? The disruption of the milk-ejection reflex is undoubtedly a common consequence of stress. Does this disruption have any adaptive benefits, or does it reflect pathological malfunction?

In fact, much less oxytocin is secreted in response to stress in lactating rats than in virgin rats. This might reflect a general hyporesponsiveness to stress during lactation because ACTH secretion in response to many stresses is also reduced. At least in part, the reduced secretion of oxytocin in lactation reflects depleted stores of oxytocin at the posterior pituitary gland because the amount of oxytocin that is secreted in response to a given activation of hypothalamic neurons is proportional to the amount of oxytocin stored in the neurosecretory nerve endings; in lactation, large amounts of oxytocin are secreted in pulses in response to suckling, and the stores are thus normally 40% or more below the levels in virgin rats. However, the pituitary stores of oxytocin increase through pregnancy, so the reduced oxytocin secretory responses to stress seen in pregnant rats requires a different explanation. The hyporesponsiveness of pregnancy seems to suggest that a specific mechanism is activated to reduce the impact of stressors on oxytocin secretion. This might be part of the mechanisms that allow pituitary stores of oxytocin to accumulate during pregnancy, or it might be part of a generalized hyporesponsiveness of stress-related signaling that is reflected by a depression of hypothalamic-pituitary-adrenal (HPA) axis responsiveness to stressors.

Much effort has been expended attempting to find some physiological rationale for the stress-induced secretion of oxytocin. These efforts have addressed three possibilities:

1. Oxytocin secreted in response to stress acts at other targets in the periphery, where its actions are appropriate and adaptive.
2. Oxytocin secreted in response to stress participates in the regulation of ACTH secretion from the anterior pituitary gland
3. The activation of oxytocin neurons by stress releases oxytocin within the brain as well as from the pituitary gland, and this release within the brain mediates the appropriate adaptive responses to stress.

Possible Peripheral Targets for Oxytocin Released during Stress

From this, it should be clear that neither the uterus nor the mammary gland seems a promising candidate as the physiological target of oxytocin that is secreted during stress. However, receptors for oxytocin are also expressed at other peripheral sites, including the heart, the thymus, and, in the rat, the kidney. Although stress induces oxytocin secretion, it is not clear which peripheral target oxytocin acts on or what its physiological effect is.

At the rat kidney, oxytocin promotes natriuresis; this provides an apparent explanation for the elevated secretion of oxytocin observed in response to increased plasma osmotic pressure. In the rat, oxytocin is also secreted in response to gastric distention, possibly indicating that it mediates a reflex natriuresis following food intake. Similar secretion of oxytocin can be evoked by peripheral injections of cholecystokinin; cholecystokinin is a gut peptide released during feeding and acts on afferent gastric vagus nerve endings. It is difficult to envisage any adaptive value to reflex natriuresis in response to generalized stress, but it seems possible that the activation of this pathway is an incidental consequence of vagal activation. This pathway seems to be a characteristic of the rat; in humans, cholecystokinin increases the secretion of vasopressin but not oxytocin.

In the rat heart, oxytocin receptors are expressed in atrial myocites, and oxytocin acting at these receptors might participate in the release of atrial natriuretic peptide (ANP). Despite its name, the major physiological role of ANP is probably not to regulate natriuresis, although this is one of its actions; instead, ANP appears to regulate the rate and force of cardiac contractions. The actions of oxytocin reduce blood pressure and reduce cardiac output – not an obviously helpful response to an acute stress, but it might be protective against excessive cardiovascular effects of sympathoadrenal activation. These actions of oxytocin seem more consistent with a role for oxytocin in body-fluid homeostasis through coordinated actions at the heart and kidney.

The thymus is responsible for the selection of the peripheral T-cell repertoire. Oxytocin, like many neuropeptides, has been identified in the human thymus by immunoreactivity and at the transcriptional level; it is present in the thymus in surprisingly large amounts in subtypes of thymic epithelial cells. Oxytocin is co-localized with the cytokeratin network of thymic epithelial cells rather than in secretory granules, so it is not secreted, but behaves like an antigen presented at the cell surface. Thymic oxytocin

also behaves as a cryptocrine signal, interacting with receptors expressed by pre-T cells. Oxytocin receptors are expressed by cytotoxic $CD8^+$ lymphocytes, and oxytocin induces phosphorylation of focal adhesion kinase in pre-T cells. Thus oxytocin might intervene in T-cell differentiation as a neuroendocrine self-antigen and as a promoter of T-cell focal adhesion. There is little evidence that circulating oxytocin has important effects at this site, and the possible adaptive value of actions of circulating oxytocin at the thymus is unclear. No convincing case has been made so far that oxytocin secreted into the peripheral circulation in response to stress has any adaptive physiological consequences at the thymus, but there has been very little investigation of this possibility.

Possible Actions of Oxytocin at the Anterior Pituitary Gland

The secretion of oxytocin in response to acute stress has prompted speculation that it might play a role in regulating ACTH secretion from the anterior pituitary gland, and pharmacological studies have indicated that oxytocin can indeed induce ACTH secretion from dispersed rat anterior pituitary cells in synergy with CRH. The ACTH-secreting activity is apparently not mediated by oxytocin receptors because oxytocin receptor mRNA is not expressed in corticotrophs and because specific oxytocin agonists have little effect on ACTH secretion; this effect might be mediated by vasopressin receptors because it can be inhibited by a vasopressin V_1 receptor antagonist. The action of oxytocin on corticotrophs suggested by *in vitro* studies has yet to be confirmed to be physiologically relevant *in vivo* in conditions of stress.

ACTH is secreted in response to the synergistic releasing activities of CRH and vasopressin; these two peptides are co-expressed in parvocellular neurosecretory neurons of the paraventricular nucleus that project to the median eminence, where these peptides are released into the hypothalamic-hypophysial portal circulation. Oxytocin secreted from the posterior pituitary gland does not appear to have direct access to the anterior pituitary gland because few blood vessels connect the separate lobes of the pituitary gland. However, some collaterals of the axons of magnocellular neurons might terminate in the median eminence, and there are also separate parvocellular populations of oxytocin neurons in the paraventricular nucleus, some of which project to the median eminence. Physiological circumstances associated with elevated secretion of oxytocin into the circulation are not consistently associated with high rates of secretion of ACTH, but the conjunction of high CRH release with high oxytocin secretion might lead to a potentiation of ACTH secretion under some circumstances because oxytocin at high concentrations is capable of triggering ACTH secretion from corticotrophs and oxytocin appears to act synergistically with CRH. It is not clear whether these actions of oxytocin are physiological rather than purely pharmacological; oxytocin receptors appear to be expressed at relatively low levels in the anterior pituitary gland, but oxytocin at high concentrations can also act at vasopressin receptors and vasopressin receptors on corticotrophs mediate the synergistic ACTH-releasing effects of vasopressin co-secreted with CRH.

A fully convincing case has yet to be made that oxytocin, released either from magnocellular neurosecretory neurons into the circulation or from parvocellular oxytocin neurons projecting to the median eminence, is important for the physiological regulation of ACTH secretion at the level of the pituitary gland, but oxytocin released centrally does appear to be involved in regulating HPA axis activity.

Oxytocin Release within the Brain

Oxytocin does not cross the blood–brain barrier in significant quantities, and oxytocin receptors do not appear to be prominently expressed at brain sites that lack an effective blood–brain barrier. Thus, oxytocin secreted from the posterior pituitary gland into the systemic circulation does not act in the brain. However, oxytocin is present at high concentrations in the cerebrospinal fluid (CSF) and at a number of sites within the brain; this oxytocin does not derive from the blood but from oxytocin released within the brain. A lot of oxytocin is released from the soma and dendrites of magnocellular neurosecretory neurons; this release, within the supraoptic and paraventricular nuclei, generally accompanies the electrical activation of the magnocellular oxytocin cells. However, there are differences in the time course of oxytocin release centrally and peripherally, and different stimuli have differential effects on the central and peripheral release of oxytocin; these observations suggest that there are important differences in the regulation of dendritic oxytocin release and release from nerve terminals. In the rat supraoptic and paraventricular nuclei, forced swimming, shaker stress, and social defeat all induce dendritic release of oxytocin from magnocellular neurons. The role for dendritic oxytocin during stress is unclear, but endogenous oxytocin modulates the activity of supraoptic oxytocin cells and their noradrenergic inputs in response to noxious stimuli.

Many stressors induce the central release of oxytocin, producing very high extracellular concentrations in several regions of the brain including the hypothalamus (especially the supraoptic and paraventricular

nuclei), the septum, and the amygdala. Central oxytocin is thought to be an antistress and anti-anxiety factor; in rats, the central injection of an oxytocin antagonist increases basal plasma concentrations of ACTH and facilitates ACTH secretion in response to various stressors and the central injection of oxytocin attenuates glucocorticoid secretion in response to stress. Rats previously exposed to the central or peripheral administration of oxytocin are less anxious when challenged in an elevated-plus maze, a behavioral test of anxiety. Central oxytocin also attenuates stress-induced CRH mRNA expression in the paraventricular nucleus and attenuates the expression of c-*fos* mRNA in stress-activated neural circuits, including the paraventricular nucleus, the septum, and the hippocampus.

The significance of the presence of oxytocin in the CSF is not clear. Although it has been speculated that the CSF might be a medium for disseminating oxytocin, mediating brain hormone-like actions, it might be that the CSF is simply a sewer, clearing oxytocin from its localized sites of activity.

The oxytocin neurons that project centrally are mainly located in the paraventricular nucleus. There are several functionally and anatomically distinct populations of parvocellular oxytocin cells in this nucleus. One population projects to the brain stem and appears to be implicated in gastric reflexes; these neurons, like magnocellular oxytocin neurons, are activated following systemic injections of cholecystokinin, and they might be part of a reflex loop that regulates gastric motility. Another population projects to the spinal cord and, in males, appears to be involved in penile erection. Oxytocin-containing fibers are also present in many other brain areas, including the preoptic area, the ventromedial nucleus of the hypothalamus, the bed nucleus of the stria terminalis, and the hippocampus. The origin of these diffusely scattered fibers is generally thought to be a separate population or populations of parvocellular oxytocin neurons in the paraventricular nucleus, whose functions are largely unknown.

Experiments involving central injections of oxytocin or of oxytocin antagonists have led to suggestions that oxytocin influences many functions and behaviors. The strongest evidence concerns ingestive behaviors (central oxytocin inhibits food intake and sodium appetite) and reproductive behaviors. Centrally acting oxytocin has been implicated in sexual behavior in both male and female rats, in the initiation of maternal behavior after parturition in the rat, and in pair bonding, notably in the prairie vole. The central release of oxytocin is undoubtedly involved in the milk-ejection reflex: intracerebroventricular injections of 1 ng oxytocin in a lactating rat lead to a striking increase in the frequency of suckling-induced milk ejections over the following 30 min, whereas intracerebroventricular injections of selective oxytocin antagonists block this reflex. These effects probably reflect the dendritic release of oxytocin from magnocellular neurons that occurs in response to suckling. Intracerebroventricular injections of much larger amounts of oxytocin (1 μg) have many consequences – including an increased activity of the HPA axis and of the sympathoadrenal system. Such doses greatly exceed the total oxytocin content of the brain and, indeed, are close to the total amount of oxytocin stored in the pituitary gland. It is not clear where in the brain the administered oxytocin acts to produce these effects, that these effects are all mediated by specific oxytocin receptors, or that they mirror any physiological role of endogenous oxytocin. Daily repeated intracerebroventricular injections of high doses of oxytocin produce contrasting effects, including lowered blood pressure, behavioral calm, and lowered activity of the HPA axis.

Suckling itself has a calming influence on the mother and, indeed as mentioned previously, the state of slow-wave sleep observed in rats to be a prerequisite for the milk-ejection reflex might be facilitated by suckling. Whether endogenous oxytocin is involved in this is not clear.

Thus, although centrally projecting oxytocin cells are activated by a variety of stresses, there is no convincing evidence yet that this activation plays any useful or adaptive role or has any important consequences. However, oxytocin is involved in a wide range of behaviors, particularly reproductive behaviors and social behaviors, that are strongly influenced by stressors. Although it has been generally assumed that the source of the oxytocin that is involved in behaviors is the centrally projecting parvocellular oxytocin neurons, there is growing evidence that oxytocin release from the dendrites of magnocellular oxytocin neurons might be important in many behaviors, including behaviors related to stress. In the final section, therefore, we indicate the key experimental evidence that has led to a current focus of attention on dendritic oxytocin release.

Dendritic Oxytocin Release and Priming

The dendrites of magnocellular neurons are large processes, packed with neurosecretory vesicles that contain oxytocin or vasopressin, and they form a dense plexus at the ventral surface of the brain below the supraoptic nucleus. We now know that the dendrites of many different neurons are electrically active, they can conduct action potentials, and they can release many neuroactive substances. Probably only the classical prejudice that dendrites are not

electrically active but are merely passive receivers of information caused us to be slow to recognize the ventral dendritic plexus of the supraoptic nucleus for what we now know it to be – a neurosecretory gland within the brain itself that is the source of very high release levels of oxytocin and vasopressin that can then diffuse to act at distant sites within the brain.

Stimuli that result in oxytocin secretion from the pituitary gland generally also result in dendritic oxytocin release, but in many circumstances dendritic release is very delayed compared to peripheral secretion, although in a few circumstances central release can precede systemic secretion. The regulation of peripheral secretion is now very well characterized – all secretion from this site is thought to be regulated by action potential activity, depolarization of neurosecretory terminals in the pituitary gland result in calcium entry into the terminals via voltage-gated calcium channels, and this calcium entry is the key step in triggering exocytosis of the large neurosecretory vesicles in which oxytocin is packaged. However, secretion from the dendrites is not all activity-dependent in the same way. Some agents, such as the peptide α-melanocyte-stimulating hormone (α-MSH), can trigger dendritic secretion even though they cause no increase at all in the action potential activity of oxytocin cells; on the contrary, α-MSH inhibits the electrical activity of oxytocin cells and thus depresses the secretion of oxytocin into the blood. α-MSH-induced oxytocin release from dendrites appears to be the result of calcium mobilization from intracellular stores rather than the result of calcium entry into the cell through membrane channels, and other peptides, including oxytocin itself, can act similarly.

Interestingly some agents that mobilize intracellular calcium stores also induce an important change in the behavior of oxytocin cells, called priming. In addition to directly causing dendritic oxytocin release, these agents also prime a readily releasable pool of oxytocin, making it available for subsequent activity-dependent release. Thus it appears that the dendrites might contain two functional pools of neurosecretory vesicles, a readily releasable pool that is available for release in response to depolarization-induced calcium entry, as at the neurosecretory terminals, and a much larger pool of vesicles that are not released in response to action potential activity but can be released by agents that mobilize calcium from intracellular stores. Such agents can augment the readily releasable pool, making more oxytocin available for subsequent activity-dependent release, and can thus exert complex and very long-lasting effects on oxytocin cell behavior.

Oxytocin neurons express oxytocin receptors; oxytocin released from dendrites acts back on oxytocin cells to modulate their activity, including by priming activity-dependent oxytocin release, and this local feedback plays an important part in the patterning of discharge of oxytocin cells. Whether oxytocin has physiological actions on other neurons in the supraoptic or paraventricular nuclei has not been established yet; but it does not act directly on magnocellular vasopressin cells. The dendritic release of oxytocin is very substantial, and spillover of this release into the CSF probably accounts for some of the oxytocin detected there. However, other populations of oxytocin cells in the paraventricular nucleus project centrally, and these other sources probably also contribute to oxytocin concentrations in the CSF.

Conclusion

In the rat, the release of oxytocin both peripherally and centrally is influenced profoundly by stress. Stresses generally increase oxytocin secretion in the rat, but they disrupt the normal pattern of secretion, leading to a hormonal message that is less effective. Thus, the interruption of milk let-down and parturition that accompanies mild stress might be physiologically appropriate; the enhanced secretion of oxytocin *per se* has no clearly established benefit. However, there is currently speculation, based mainly on behavioral effects seen after the central administration of high doses of oxytocin, that the central release of oxytocin has a generally calming influence.

Further Reading

Amico, J. A., Mantella, R. C., Vollmer, R. R., et al. (2004). Anxiety and stress responses in female oxytocin deficient mice. *Journal of Neuroendocrinology* **16**, 319–324.

Douglas, A. J. (2005). Central noradrenergic mechanisms underlying acute stress responses of the hypothalamo-pituitary-adrenal axis: adaptations through pregnancy and lactation. *Stress* **8**, 5–18.

Geenen, V., Keeha, O. and Martens, H. (1998). Thymic expression of neuroendocrine self-peptide precursors: role in T cell survival and self-tolerance. *Journal of Neuroendocrinology* **10**, 811–822.

Jankowski, M., Hajjar, F., AlKawas, S., et al. (1998). Rat heart: a site of oxytocin production and action. *Proceedings of the National Academy of Sciences USA* **95**, 14558–14563.

Leng, G. and Brown, D. (1997). The origins and significance of pulsatility in hormone secretion from the pituitary. *Journal of Neuroendocrinology* **9**, 493–513.

Ludwig, M. (1998). Dendritic release of vasopressin and oxytocin. *Journal of Neuroendocrinology* **10**, 881–895.

Ludwig, M., Sabatier, N., Bull, P. M., et al. (2002). Intracellular calcium stores regulate activity-dependent neuropeptide release from dendrites. *Nature* **418**, 85–89.

Onaka, T. (2004). Neural pathways controlling central and peripheral oxytocin release during stress. *Journal of Neuroendocrinology* **16**, 308–312.

Russell, J. A., Leng, G. and Douglas, A. J. (2003). The magnocellular oxytocin system, the fount of maternity: adaptations in pregnancy. *Frontiers in Neuroendocrinology* **24**, 27–61.

Sabatier, N., Caquineau, C., Bull, P., et al. (2003). Alpha-melanocyte-stimulating hormone stimulates oxytocin release from the dendrites of hypothalamic neurons while inhibiting oxytocin release from their terminals in the neurohypophysis. *Journal of Neuroscience* **23**, 10351–10358.

Stricker, E. M. and Verbalis, J. G. (1996). Central inhibition of salt appetite by oxytocin in rats. *Regulatory Peptides* **66**, 83–85.

Uvnas-Moberg, K. (1998). Antistress pattern induced by oxytocin. *News in Physiological Sciences* **13**, 22–26.

P

Pain

H J Strausbaugh and J D Levine
University of California, San Francisco, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 115–118, © 2000, Elsevier Inc.

Introduction
Stress-Induced Analgesia
Stress-Induced Hyperalgesia

Glossary

Analgesia	A decreased perception of pain.
Hyperalgesia	An increased perception of pain.
Nociception	A perception of pain.
Opioid	Any of three families of endogenous peptides (i.e., as endorphin, enkephalin, and dynorphin) that induce analgesia by binding to μ, δ, or κ opioid receptors.

Introduction

Definition and Measurement of Pain

Pain has been defined as an aversive sensation originating from a noxious stimulus in a defined area of the body. Such intense, potentially tissue-injurious stimuli activate nociceptive sensory neurons (nociceptors), which transmit signals via the spinal cord to the brain that are perceived by the organism as pain. The central nociceptive processing areas include the thalamus, hypothalamus, mesencephalic tegmental area, mesolimbic areas, parabrachial nuclei, amygdala, bulbopontine reticular formation, and the somatosensory cortex. Pain has two components: a sensory or discriminate component and an affective or emotional component. Because pain involves both of these components, simply recording from peripheral nociceptors does not provide a sufficient measurement of pain because the affective component of pain is not assessed in this manner. To circumvent this problem, behavioral measures of pain are used. If animals display characteristic behaviors after application of a noxious stimulus, including withdrawal of the area to which the stimulus is applied, licking of the affected area, vocalization, or escape behavior, it is inferred that the animal is experiencing pain. A commonly used behavioral test for the measurement of pain in animals and the test upon which much of the current information on stress effects on pain is based is the tail flick test. In this test, an animal is lightly restrained and its tail is exposed to a heat source intense enough to produce pain in humans. The time it takes for the animal to remove its tail from the heat source is recorded as tail flick latency; the more painful the stimulus, the shorter the latency. An additional tool used to identify pain is the administration of opioid analgesics such as morphine. Morphine activates an endogenous pain control circuit to block the transmission of nociceptive signals. Therefore, if a noxious stimulus-induced behavior is blocked by morphine, it is considered to be painful.

Definition and Measurement of Stress

Stress is generally defined as any stimulus to which the organism is not adapted. Like pain, stress involves a perceptual component and is therefore difficult to assess in laboratory animals. To circumvent this problem, hormonal measurements are commonly used. Stressful stimuli activate three neuroendocrine circuits: the hypothalamic-pituitary-adrenocortical (HPA) axis, the sympathoadrenal axis, and the sympathetic neural axis. Although different stressors may activate these axes to varying degrees, all axes are activated by most stressful stimuli. Therefore, the hormonal measurement of activation of one of these axes, generally the HPA axis, is used as an operational definition of stress. Both adrenocorticotropic hormone (ACTH) and corticosterone are secreted rapidly into the blood upon activation of the HPA axis. Stimuli that evoke increases beyond normal circadian

levels in either of these hormones are considered stressful.

How do these broad, somewhat overlapping concepts, pain and stress, relate to each other? One of the first written observations of the relationship between pain and stress comes from observations made by field doctors in Europe during World War II. These doctors observed that up to 70% of soldiers presenting with severe battle wounds would not report significant levels of pain. This apparent insensitivity to pain was short-lived; 24 h later, all soldiers reported significant levels of pain. Reports such as these suggested that the perception of pain could be modified cognitively. Because stress seemed to be a common factor in all of these reports, the question arose as to whether stress could affect pain perception. In 1976, a decreased perception of pain after exposure to a stressor was shown in laboratory experiments for the first time and was termed stress-induced analgesia. Although less well studied, it has also been shown that stress can enhance pain perception in some circumstances. This phenomenon has been termed stress hyperalgesia and is generally referred to as stress-induced hyperalgesia.

Stress-Induced Analgesia

Stress-induced analgesia refers to the well-established phenomenon that exposing humans or animals to a wide variety of stressors induces the suppression of pain perception. Although electric foot shock and cold water swim have been the most well-studied eliciting stressors, others include rotation, immobilization, and food deprivation. The decrease in pain sensitivity in these studies is usually measured as an increase in tail flick latency to noxious thermal stimuli. However, other pain sensitivity tests such as paw withdrawal, paw licking, and escape behavior in response to exposure to electrified grids or hot plates are also used.

Although the mechanisms of stress-induced analgesia are not completely understood, two subtypes of stress-induced analgesia – endogenous opioid dependent (opioid mediated) and opioid independent (nonopioid mediated) – can be distinguished. Opioid-mediated stress-induced analgesia is defined as stress-induced analgesia that either is blocked by the opiate receptor antagonist naloxone or is cross-tolerant to opioid agonists such as morphine. Cross-tolerance is determined by treating an animal for several days with morphine until the drug no longer produces analgesia. A stressful stimulus is then applied, and if the analgesia is blocked, then it is considered cross-tolerant to morphine. The nonopioid-mediated type of stress-induced analgesia is unaffected by these treatments. Characteristics of the stressor, including the temporal pattern of administration and severity, seem to predict whether opioid or nonopioid analgesia will be induced. In general, stressors that are intermittent and less severe produce opioid-mediated analgesia, whereas stressors that are continuous and of greater severity produce nonopioid-mediated analgesia. Both forms can be classically conditioned.

The distinct nature of opioid- and nonopioid-mediated subtypes of stress-induced analgesia is supported by the fact that they are not cross-tolerant. In fact, a collateral inhibition model has been proposed by Bodnar and colleagues in which activation of one form of stress-induced analgesia would actively inhibit the other form. These investigators have shown that, if administered closely in time, nonopioid-mediated analgesia inhibits opioid-mediated analgesia. This is in contrast to the well-established phenomenon that distinct elicitors of opioid-mediated stress-induced analgesia synergize to produce greater analgesia. Bodnar and colleagues propose that collateral inhibition allows the animal to use its most adaptive system in response to potentially widely varying stressful situations presented by the environment.

The mechanism of opioid-mediated stress-induced analgesia is thought to involve both central and peripheral sources of opioids as well as both central and peripheral sites of action. Stress activates endogenous opioidergic neural circuits in the midbrain, which then transmit signals through the dorsolateral funiculus to inhibit pain transmission circuits in the dorsal horn of the spinal cord. Stress also induces the release of opioids from peripheral sources, including the pituitary gland and the adrenal medulla, and, during inflammation, from leukocytes. These opioids may act by binding to receptors on terminals of nociceptive peripheral sensory nerves and inhibiting noxious stimulation-induced activation or sensitization of these neurons. It is important to note that while these opioidergic pathways are a key element in mediating this type of analgesia, other neurotransmitters (e.g., serotonin, norepinephrine, acetylcholine), neuropeptides (e.g., cholecystokinin), and hormones (e.g., corticosterone) may also participate in this pathway.

Although it is clear that distinct pathways mediate opioid and nonopioid forms of stress-induced analgesia, there is some evidence that there may also be some pathways in common. Similar to opioid-mediated forms of stress-induced analgesia, both midbrain structures and the dorsolateral funiculus of the spinal cord are involved in mediating nonopioid stress-induced forms of analgesia. Additionally, mouse strains have been developed that express high or low levels of

stress-induced analgesia. This phenotype is present whether the animals are exposed to stressors that characteristically induce opioid- or nonopioid-mediated analgesia. However, in the absence of specific gene products, this is not strong evidence for an overlap in the mechanism but does offer a promising avenue for future studies. In contrast to opioid-mediated stress-induced analgesia, vasopressin, thyrotropin-releasing hormone, and histaminergic systems are involved in mediating nonopioid-mediated stress-induced an algesia. Some forms of nonopioid-mediated stress-induced analgesia (e.g., continuous cold water swim) are mediated by the HPA axis and, in particular, by the adrenocortical system, whereas other forms (e.g., continuous foot shock of short duration) are not. Additionally, dopaminergic and cholinergic systems may play a role in mediating continuous cold water swim analgesia. Clearly, further research is required to elucidate the pathways that mediate nonopioid-mediated stress-induced analgesia.

Stress-Induced Hyperalgesia

Stress-induced hyperalgesia is defined as an increased sensitivity to noxious stimuli or a decrease in analgesia after exposure to stress. In the laboratory, an increased sensitivity to noxious stimuli is usually indicated by a decreased tail flick latency to noxious heat stimulation. Stress-induced hyperalgesia has been observed in animals after exposure to a variety of stressors, including ether, mechanical oscillation, vibration, a novel environment, and repeated exposure to cold temperatures. Stress-induced hyperalgesia has been much less well studied than stress-induced analgesia, and consequently the mechanisms of stress-induced hyperalgesia are still poorly understood. However, a general concept is emerging that an induction of stress-induced hyperalgesia is due to changes in the central, possibly perceptual, processing of nociceptive information by the organism and not due to changes in the processing of noxious stimuli at the level of the peripheral nociceptive pathways as may contribute to stress-induced analgesia. This emerging concept is based on several lines of evidence. First, electrophysiological recordings from nociceptive sensory neurons in the peripheral nerves indicate no change in sensitivity to the noxious stimulus before and after stress-induced hyperalgesia is elicited. Second, stress-induced hyperalgesia can be classically conditioned such that even when a noxious stimulus is not applied, a neutral stimulus that had been paired previously with the noxious stimulus can elicit hyperalgesia. Third, emotional characteristics of the rat can determine whether stress-induced hyperalgesia is induced. Although it is difficult to assess the emotional state of a rat, studies have shown that rats displaying behaviors thought to indicate anxiety (e.g., vocalization, agitation) during exposure to stress and the noxious stimulus display stress-induced hyperalgesia, whereas rats not displaying these behaviors and exposed to identical stimuli do not show stress-induced hyperalgesia. This concept is supported by the fact that administering antianxiety drugs (e.g., clonidine, diazepam) to these rats blocks stress-induced hyperalgesia, whereas administering anxiety-producing drugs (e.g., yohimbine) enhances and prolongs stress-induced hyperalgesia. Finally, clinical studies indicate that anxious patients display an increased sensitivity to pain. Of course, in these studies it is difficult to assess whether patients are more sensitive to pain because they are anxious or whether they are anxious because they are perceiving more pain. Further investigation will be required to elucidate the factors that elicit and the mechanisms that produce stress-induced hyperalgesia. Additionally, it has been shown that stress may inhibit the analgesic effects of drugs as well as enhance pain perception; however, the relationship between these two effects is currently unknown. Further investigation will be required to distinguish these two phenomena.

See Also the Following Articles

Hypothalamic-Pituitary-Adrenal; Opioids.

Further Reading

Amit, Z. and Galina, Z. H. (1986). Stress-induced analgesia: adaptive pain suppression. *Physiological Reviews* **66**, 1091–1120.

Bardiani, A. and Pavone, F. (1991). Reduction of oxotremorine-induced analgesia after chronic but not acute restraint stress. *Psychopharmacology (Berlin)* **104**, 57–61.

Bodnar, R. J. (1990). Effects of opioid peptides on peripheral stimulation and "stress"-induced analgesia in animals. *Critical Reviews in Neurobiology* **6**, 39–49.

Foa, E. B., Zinbarg, R. and Rothbaum, B. O. (1992). Uncontrollability and unpredictability in post-traumatic stress disorder: an initial model. *Psychological Bulletin* **112**, 218–238.

Gamaro, G. D., Xavier, M. H., Denardin, J. D., et al. (1998). The effects of acute and repeated restraint stress on the nociceptive response in rats. *Physiology and Behavior* **63**, 693–697.

Gue, M., Del Rio-Lacheze, C., Eutamene, H., Theodorou, V., Fioramonti, J. and Bueno, L. (1997). Stress-induced visceral hypersensitivity to rectal distension in rats: role

of CRF and mast cells. *Neurogastroenterology and Motility* **9**, 271–279.
Kelly, D. D. (ed.) (1986). *Stress-induced analgesia*. Annals of the New York Academy of Sciences (Vol. 467). New York: New York Academy of Sciences.
King, T. E., Crown, E. D., Sieve, A. N., Joynes, R. L., Grau, J. W. and Meagher, M. W. (1999). Shock-induced hyperalgesia: evidence forebrain systems play an essential role. *Behavioural Brain Research* **100**, 33–42.
Madden, J., Akil, H., Patrick, R. L. and Barchas, J. D. (1977). Stress-induced parallel changes in central opioid levels and pain responsiveness in the rat. *Nature* **265**, 358–360.
Mogil, J. S., Sternberg, W. F., Marek, P., Sadowski, B., Belknap, J. K. and Liebeskind, J. C. (1996). The genetics of pain and pain inhibition. *Proceedings of the National Academy of Sciences USA* **93**, 3048–3055.
Murison, R. and Overmier, J. B. (1993). Parallelism among stress effects on ulcer, immunosuppression and analgesia: commonality of mechanisms? *Journal of Physiology (Paris)* **87**, 253–259.
Nijenhuis, E. R., Vanderlinden, J. and Spinhoven, P. (1998). Animal defensive reactions as a model for trauma-induced dissociative reactions. *Journal of Trauma and Stress* **11**, 243–260.
Simone, D. A. (1992). Neural mechanisms of hyperalgesia. *Current Opinion in Neurobiology* **2**, 479–483.
Vidal, C. and Jacob, J. J. (1982). Stress hyperalgesia in rats: an experimental animal model of anxiogenic hyperalgesia in human. *Life Sciences* **31**, 1241–1244.
Watkins, L. R. and Mayer, D. J. (1982). Organization of endogenous opiate and nonopiate pain control systems. *Science* **216**, 1185–1192.

Panic Disorder and Agoraphobia

J C Ballenger
Medical University of South Carolina, Charleston, SC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J C Ballenger, volume 3, pp 119–123, © 2000, Elsevier Inc.

Symptoms
Diagnosis
PD in the General Medical Setting
Epidemiology
Risk Factors
Course and Prognosis
Treatment
Psychological Treatments: Psychodynamic Psychotherapy
Behavioral Treatments
Selection of Treatments

Glossary

Panic attacks (PA)	The name used currently for unexpected attacks of fear.
Panic disorder (PD)	Originates from the Greek god of flocks, Pan, who would frighten animals and humans out of the blue.
Spontaneous panic attacks	The main characteristic of PD that is central to its recognition and diagnosis.

Symptoms

One of the main characteristics of panic disorder (PD) is spontaneous out of the blue panic attacks (PAs). PAs usually occur immediately upon exposure or in anticipation of exposure to particular situations, often ones in which panic attacks have previously occurred.

Symptoms of PAs in order of their frequency are palpitations, pounding heart, tachycardia, sweating, trembling or shaking, shortness of breath or smothering, feeling of choking, chest pain or discomfort, nausea or abdominal distress, feeling dizzy, unsteady, lightheaded or faint, derealization or depersonalization, fear of losing control or going crazy, fear of dying, chills, and hot flashes.

Panic attacks usually last for several minutes, but they can also last for hours. In one of the largest modern PD clinical trials, patients averaged one to two PAs per week. However, the frequency and severity of PAs vary greatly between individuals and, at times, in individuals. Some may have only one to three PAs per year, whereas others may have multiple PAs on a daily basis. Some individuals will even have bursts of PAs and then not experience attacks for a long period of time.

Because PAs are extremely frightening, patients will often develop a logical fear that they will reoccur. This anticipatory anxiety increases prior to exposure to situations in which patients have experienced previous PAs. Because of this, PD patients often develop

agoraphobia, which is avoidance of certain places or situations. Agoraphobics often avoid situations in which they have had PAs because they are afraid that escape would be difficult or embarrassing or that help might not be available during these attacks. In community samples, one-third to one-half of patients who meet criteria for PD also have significant agoraphobia avoidance.

There are typical situations that are avoided most frequently. These include taking buses, trains, or planes; riding in or driving a car; being in large crowds; standing in lines; shopping; and being in spaces where they might feel trapped (e.g., bridges, tunnels, elevators). Patients experiencing a PA will often have an overwhelming need to escape or return to a place of safety, like home. These patients do not actually fear the situation itself but are afraid of the feelings of panic that might happen while in that situation.

Many patients even have PAs that begin during sleep. These nocturnal panic attacks are quite common, and the majority of PD patients experience them. These PAs occur during the beginning states of the sleep cycle and are essentially identical to daytime PAs.

There also is a close relationship between depression and PD. In various samples, comorbid or secondary depression ranged from 22 to 68%. Proper recognition of comorbid depression is especially important because of the marked increase in suicide attempts (23.6–50%) in patients with both PD and depression.

Diagnosis

The diagnosis of PD has several requirements.

1. Recurrent, unexpected PAs (situational PAs could also occur, but there would need to be at least two unexpected PAs by history).
2. Panic attacks need to be followed by at least 1 month of (1) persistent anxiety about the potential recurrence of further PAs, (2) persistent anxiety about the implications of these attacks (e.g., going crazy, something wrong medically), or (3) a significant behavioral change because of these attacks.

PD in the General Medical Setting

There is poor recognition of PD in the general medical setting. It is often undiagnosed, because physical symptoms of PD can be mistakenly linked to other medical conditions. However, panic-like symptoms can occur in a few medical conditions (e.g., hyperthyroidism, pheochromocytoma, seizures, cardiac arrhythmias, and chronic obstructive pulmonary disease).

Conservative estimates of PD in primary care have ranged from 3 to 8%, with at least 50% of patients with PD being unrecognized. The average length of time for a patient in the general medicine setting to be diagnosed correctly with PD is 10 years, with an escalation of the use of health-care services over this 10-year period. Overall, PD patients are three times more likely to utilize general medical services than the traditional patient in a primary care setting. The percentage of PD patients is also markedly higher in certain medical groups (e.g., cardiology and gastroenterology).

A study sponsored by the World Health Organization (WHO) studied primary care patients in 14 different countries. Ninety-nine percent of patients with one PA in the previous month had an anxiety disorder or depression or a subthreshold anxiety disorder or depressive disorder. If replicated, it would appear that the occurrence of even a single PA should serve as a signal for increased scrutiny for anxiety and depressive syndromes.

Epidemiology

There is a striking uniformity worldwide for the observed prevalence of PD, generally averaging between 1 and 2% in multiple studies. In the WHO Primary Care Survey of 14 countries mentioned previously, the prevalence for PD ranged from a low of 1.4% to a high of 16.5% for PAs and 0–03.5% for PD itself. The average for PD was 1.1% (current) and 3.5% (lifetime), which was similar to community samples. As mentioned, rates are much higher in certain medical settings, ranging from 15% in dizziness clinics to 16–65% in cardiology practices and 35% in hyperventilation clinics.

Risk Factors

In treatment samples, PD has been observed to be at least two times more prevalent in females than in males. Males tend to have a longer duration of illness but are less likely to seek help or develop agoraphobia and depression. The onset of PD for both genders usually occurs between 15 and 24 years of age, with a second peak at 45–54 years of age. The onset of PD after the age of 65 is rare (0.1%).

As mentioned previously, agoraphobia is also more prevalent in females, with the highest rates of PD and agoraphobia occurring in widowed, divorced, or separated individuals living in cities. Early parental loss, limited education, and physical or sexual abuse are also risk factors for PD.

Genetic factors can even be associated with PD. Heightened anxiety, known as behavioral inhibition,

is higher in children of anxiety-disordered parents and in children of adults with PD. As these children have aged, they have been observed to have higher rates of anxiety and phobic disorders.

Precipitating events have been reported in 60–96% of PD cases. Multiple studies also suggest that traumatic early events may figure into the vulnerability leading to PD.

Course and Prognosis

Available evidence suggests that PD is a chronic condition, once criteria for the disorder are met. It is common for patients to report symptoms that they have had for 10–15 years prior to diagnosis. As mentioned previously, the Klein model suggests that spontaneous PAs are the first manifestation of the illness, followed by anticipatory anxiety and then agoraphobia in some individuals.

In fact, most PD patients do not have PAs as their first symptoms. Over 90% have had mild phobic or hypochondriac symptoms prior to the onset of their first PA. The first PA is usually experienced in a phobogenic situation, such as a public place, street, public transportation, crowd, elevator, tunnel, bridge, or open space. Despite this, the Klein model remains largely correct and helpful in conceptualizing the development of the condition and its treatments.

After diagnosis and some sort of treatment, functional recovery occurs in the majority of patients. Poor responses to treatment were associated most consistently with initial high symptom severity and high agoraphobic avoidance at baseline. A poor response is also associated with a low socioeconomic status, less education, longer duration, limited social networks, death of a parent, divorce or unmarried status, and personality disorders.

Treatment

Since the early 1960s, effective psychological and pharmacological treatments for PD have been developed.

Tricyclic Antidepressants (TCAs)

In the 1960s, the TCA imipramine was one of the first effective drug treatments. Important advantages of imipramine include its once-a-day dose, its low cost, and the extensive research that has been conducted on its effectiveness. Its principal difficulties are initial hyperstimulation reactions, significant weight gain with long-term use, and the dangers involved with an overdose. The highly serotonergic clomipramine became the most popular treatment in Europe during the 1970s, 1980s, and 1990s.

Monoamine Oxidase Inhibitors (MAOIs)

During the 1960s, MAOIs were also proven to be effective in the treatment of PD. The most definitive trial was an early comparison of phenelzine and imipramine in the mid-1970s. This was the first trial to suggest that MAOIs might be the most effective medication class, especially if the PD patient is depressed. The biggest drawbacks of MAOIs are the need for a patient to be on a restricted diet and the danger of hypertensive crisis if the diet is not followed. There are also dangers if certain other medications are used, especially certain analgesics and other antidepressants. The principal side effects are insomnia, weight gain, postural hypotension, anticholinergic side effects, and sexual dysfunction.

The reversible MAOIs, brofaramine and meclobemide, have stimulated interest because they require no restrictive diet. However, trials with meclobemide were unable to distinguish medication from placebo, and despite promising early results with brofaramine, its development has been discontinued. There is still active research in this area.

Benzodiazepines (BZs)

High potency BZs (e.g., alprazolam, clonazepam) were developed in the 1980s and are also used in the treatment of PD. In fact, BZs remain the most frequently utilized treatment for PD worldwide, probably because of their quick onset of action, increased efficacy against anticipatory anxiety, low cost, and tolerability. They are frequently utilized in conjunction with the antidepressants. The principal drawback to BZs is the general concern with their abuse potential, although it is alcoholics and opiate or pill addicts that abuse the drug. Abuse in uncomplicated PD patients is extremely rare. Other side effects include the high rate of withdrawal symptoms and relapse that can occur after discontinuing BZ treatment.

Selective Serotonin Reuptake Inhibitors (SSRIs)

More recent studies in the 1990s have demonstrated that the SSRIs are effective, and they are now frequently recommended as the first treatment of choice. This is largely based on the many controlled trials documenting the high tolerability of SSRIs and the decreased rate of jitteriness and hyperstimulation responses often associated with tricyclic antidepressants. They are preferred over benzodiazepines in patients with significant difficulty with depression. Most (70–85%) patients retain their benefit when SSRIs are tapered and discontinued after 12 months of successful treatment.

Psychological Treatments: Psychodynamic Psychotherapy

Although psychodynamic psychotherapy remains a popular treatment for many psychiatric disorders, including PD, there is limited research demonstrating its efficacy in PD. An emotion-focused treatment for PD has been developed that explores typical fears of being abandoned or trapped as stimuli for panic attacks. This often involves a 12-session acute treatment with 6 sessions of monthly maintenance. Patients are encouraged to identify, reflect upon, and attempt to change problematic feelings and their responses.

Behavioral Treatments

There is a rapidly increasing body of evidence for combinations of exposure-based treatments and cognitive-behavioral therapy (CBT) that are effective in the treatment of PD.

Exposure Treatments

Behavioral treatments, involving *in vivo* exposure, require patients to gradually systematically re-expose themselves to their phobic situations. A large number of studies have documented that 58–83% of patients improve after exposure treatments.

One consistent observation of exposure-based treatments is its long-lasting effects. In one trial with 110 individuals treated with 12 weeks of exposure-based treatment, 74% of the sample achieved panic-free status, and this improvement was maintained in almost all patients throughout a 7-year follow-up. Studies that have compared the use of exposure to CBT have found them to be essentially equal in efficacy. There appears to be little benefit from the combination, although the dropout rate was lower.

Evidence supporting the importance of exposure-based treatments is very clear, but many patients remain symptomatic. In addition, many studies report that between 10 and 25% of patients drop out of the therapy because it is an anxiety-provoking treatment method.

Cognitive-Behavioral Therapy

Cognitive-behavioral therapy is based on the theory that patients with PD interpret physical symptoms in a catastrophic way and that these cognitive distortions need to be challenged and corrected. CBT of PD evolved from early work of Aaron Beck but has been applied to PD primarily by Barlow and colleagues and is called panic control therapy (PCT).

Treatment generally involves an initial education component, which is followed by the identification of critical misinterpretations of panic symptoms. Patients are then taught ways in which they can challenge and correct these misinterpretations. This includes evaluating how likely (or unlikely) the imagined consequences are and how catastrophic it would be even if their worse fears did happen. Most CBT is also combined with interoceptive exposure to the physical symptoms that frighten them. For instance, a patient fearful of dizziness might be spun in a chair.

Although applied relaxation has been shown to be somewhat effective in PD, other studies have demonstrated that CBT, especially combined with interoceptive exposure (i.e., PCT), is more effective than medication.

Most experts would agree that the treatment of PD with pharmacotherapy and a CBT approach – employing a combination of exposure and cognitive therapy – are roughly comparable in efficacy. The most carefully performed trial demonstrating this was a multicenter, 11-week acute trial comparing imipramine to CBT. Both CBT and imipramine performed equally well, but the combination of CBT and imipramine was not proven to be any more effective in the treatment of PD. This trial provided evidence that CBT applied early in treatment did not offer protective effects against imipramine relapse. However, other data suggest that it is protective if given during the discontinuation period.

Selection of Treatments

Pharmacotherapy (BZs or antidepressants) and CBT/exposure treatments appear to be equally effective, although response is perhaps faster with medication treatment and longer term response is better in some patients with CBT/exposure-based treatments. Most often, selection of treatment is determined by patient choice or individual characteristics.

One important issue in the treatment of PD is that the majority of patients do not fully recover with either pharmacotherapy or psychotherapy. For both types of treatment, only 25–45% of patients could be considered fully recovered, despite the excellent progress that has been made since the 1980s. This has led to ongoing research attempts to find more effective treatment methods for PD.

See Also the Following Articles

Anxiety; Cognitive Behavioral Therapy; Pharmacological Treatments of Stress.

Further Reading

Ballenger, J. C. (1998). Treatment of panic disorder in the general medical setting. *Journal of Psychosomatic Research* **44**, 5–15.

Ballenger, J. C. (1998). Comorbidity of panic and depression: Implications for clinical management. *International Journal of Clinical Psychopharmacology* **13**(Supplement), 513–517.

Ballenger, J. C. (2003). Selective serotonin reuptake inhibitors in the treatment of the anxiety disorders. In: Nutt, D. & Ballenger, J. C. (eds.) *Anxiety disorders*, pp. 339–361. Oxford, UK: Blackwell Science.

Ballenger, J. C., Burrows, G. and DuPont, R. L., Jr., et al. (1988). Alprazolam in panic disorder and agoraphobia. Results from a multicenter trial. I. Efficacy in short-term treatment. *Archives of General Psychiatry* **455**, 413–422.

Ballenger, J. C., Lydiard, R. B. and Turner, S. M. (1997). Panic disorder and agoraphobia. In: Gabbard, G. O. (ed.) *Treatments of psychiatric disorders* (2nd edn., vol. 2), pp. 1421–1452. Washington, D.C.: American Psychiatric Press.

Ballenger, J. C., Davidson, J. R. T., Lecrubier, Y. and Nutt, D. J. (International Consensus Group on Depression and Anxiety), et al. (1998). Consensus statement on panic disorder from the International Consensus Group on Depression and Anxiety. *Journal of Clinical Psychiatry* **59**, 47–54.

Klein, D. F. (1964). Delineation of two drug responsive anxiety syndromes. *Psychopharmacology* **5**, 397–408.

Klerman, G. L., Weissman, M., Ovellette, R., Johnson, J. and Greenwald, S. (1991). Panic attacks in the community: social morbidity and health care utilization. *Journal of the American Medical Association* **265**, 742–746.

Margraf, G., Barlow, D. H., Clark, D. M. and Telch, M. J. (1993). Psychological treatment of panic: work in progress in outcome, active ingredients, and follow-up. *Behavioral Research Therapy* **31**, 1–8.

Marks, I. M. (1969). *Fears and phobias*. London: Heinemann.

Mavissakalian, M. R. and Prien, R. F. (eds.) (1996). *Long-term treatments of anxiety disorders*. Washington, D.C.: American Psychiatric Press.

Rosenbaum, J. F., Biederman, J., Gersten, M., et al. (1988). Behavioral inhibition in children of parents with panic disorder and agoraphobia: a controlled study. *Archives of General Psychiatry* **45**, 463–470.

Roy-Byrne, P. P. and Cowley, D. S. (1995). Course and outcome in panic disorder. A review of recent follow-up studies. *Anxiety* **1**, 151–160.

Sartorius, N., Ustun, T. B., Jorges-Alberto, C., et al. (1993). An international study of psychological problems in primary care. *Archives of General Psychiatry* **50**, 819–824.

Paranoia

P Kinderman
University of Liverpool, Liverpool, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Kinderman, volume 3, pp 124–129, © 2000, Elsevier Inc.

Definition and Description
Significance and Social Impact of Paranoia
Assessment
Theories of Paranoia
Treatment of Paranoia

Glossary

Causal attribution	An explanation for the occurrence of events.
Delusion	A false personal belief sustained in spite of apparently incontrovertible and obvious proof or evidence to the contrary.
Paranoia	An unreasonable belief held by a person that other people are victimizing him or her; an unreasonable fear of threat.
Paranoid personality disorder	A disorder characterized by a pervasive distrust and suspiciousness of others, such that their motives are interpreted as malevolent.

Definition and Description

The meaning of paranoia and paranoid has fluctuated over the centuries. The literal translation of the Greek paranoid is "by the side of the mind" and was used in a general manner to mean crazy or mad. The use of the term paranoid subsided until its resurgence in Germany in the first half of the nineteenth century to refer to a distinct delusional disorder. In the English-speaking world, a distinct confusion of use emerged, with the United Kingdom tending to mirror continental Europe and using the word paranoia to describe systematized delusional insanity. In the United States, the acceptance of Freud's 1911 psychodynamic model

of paranoia led to the widespread use of the term and to a diminution of any distinctions between delusional and nondelusional paranoia.

The links among paranoia, persecutory delusions, and paranoid schizophrenia remain close but confused. In ordinary language, however, the word paranoid has become general, applied to people who either have unreasonable or unwarranted complaints of victimization unreasonable fears of threat.

At its extreme, such concerns may lead to the label of paranoid personality disorder, defined as a pervasive distrust and suspiciousness of others such that their motives are interpreted as malevolent. People with this disorder are described as being suspicious, being cynical, and tending to bear grudges and take criticism extremely badly. Paranoid personality disorder appears to be relatively common in the general population and has been placed on a continuum in which paranoid schizophrenia is seen as a more severe form of paranoia, which is in turn seen as a more severe form of paranoid personality disorder.

Significance and Social Impact of Paranoia

The terms paranoid or paranoia are used repeatedly in political and social contexts. Paranoia and persecutory delusions are also commonly associated with violence. Paranoia has been described as a characteristic of a vast variety of political systems. The societies in Plato's *Republic* and Hobbes's *Leviathan* have been described as characteristically paranoid as have a great number of political systems from Tudor England between the fifteenth and seventeenth centuries, the United Kingdom in the eighteenth to nineteenth centuries, and both the U.S. and Soviet Cold War political systems. Political leaders as individuals (particularly Hitler, Stalin, and Nixon) have also been described as paranoid.

Paranoid delusions have been associated with many violent acts. The M'Naughten rules, the legal basis for an insanity defense, were established after Daniel M'Naughten attempted to kill British Prime Minister Sir Robert Peel in 1843. M'Naughten mistook the prime minister's aide, Edward Drummond, for Sir Robert and shot him instead. M'Naughten was found to be clearly and delusionally paranoid. According to an analysis of the so-called White House cases (individuals who were hospitalized because of their psychotic preoccupation with U.S. politicians), paranoid delusions were common, but the best predictor of arrest for a violent crime after hospital discharge was an arrest for a violent crime before hospitalization. This finding has been generally supported; the likelihood of acting on delusions is generally unrelated to phenomenology of the delusion but is predicted by the perceived evidential support for the belief, negative emotions, and social functioning.

Assessment

Paranoid ideation can be assessed in its own right as delusional beliefs or as a symptom of schizophrenia.

Delusional classification is complex, and it is often difficult to determine the difference between unusual beliefs and full delusions, between brief episodes and persistent delusions, or between bizarre versus nonbizarre delusions. Thus, continua rather than dichotomous classifications may be appropriate. Delusions are also usually assessed on a variety of parameters: conviction, cultural or stimulus determinants, preoccupation, implausibility, extension (the degree to which the delusional belief involves various areas of the patient's life), bizarreness, disorganization, and emotional commitment.

General psychiatric rating scales include brief assessments of the presence or absence of the symptoms of schizophrenia, including delusional beliefs. A more precise assessment can be obtained using specific research devices. Within such scales, however, delusional ideation is viewed as a part of a whole syndrome rather than as a phenomenon worth examining alone. More specific rating scales and assessment devices have been developed for the assessment of delusions alone.

Despite doubts about the validity and reliability of scales for the assessment of paranoia, the fact that they seem incestuous in their development, and their reliance on doubtful theories of personality structure, diagnosed patients with delusions of persecution score high on these measures, whereas nonpatient samples show predictable associations with other measures of cognitive functioning. However, paranoid ideation is complex. It may be the case that measures applicable for normal and clinical populations are different.

Theories of Paranoia

Theories of paranoia are occasionally subsumed under theories of the development of delusions or schizophrenia. The presumed causes of schizophrenia are not dealt with here; note, however, that nearly all variables believed to influence human behavior have been invoked. Specific theories concerning the etiology of delusions are equally numerous.

Biological Approaches

Medical and psychiatric conditions Paranoid delusions occur in association with at least 105 general

medical conditions. In many nonpsychiatric conditions associated with delusions, delirium (a state of disordered or impaired consciousness combined with an alteration of perceptual functioning or affect) seems important. In such medical conditions, paranoid ideas are presumed to be secondary to the physical pathology. Although it seems difficult to understand the etiology of the abnormal beliefs, it does seem that when the underlying illness is treated the delusions usually disappear. Paranoid delusions are far more common in psychiatric conditions, in which they are often difficult to treat and tend to persist.

Delusions occur in 10–20% of depressed inpatients, with mood-congruent delusions being more common. Delusions occur in up to 70% of manic patients and are usually grandiose, concerning wealth, importance, status, or special powers. As in depression, separating mood-congruent, exaggerated, and abnormal (but nondelusional) beliefs from true delusions in mania is difficult and perhaps of little clinical value.

The presence of delusions is highly likely to lead to a diagnosis of paranoid schizophrenia. Delusions of persecution have been reported in over 60% of patients with a diagnosis of paranoid schizophrenia. Such delusions in paranoid schizophrenia are usually systematized and persistent, as opposed to the fleeting and haphazard beliefs of people with organic diseases.

Genetics Genetic studies of paranoid schizophrenia seem to imply that inherited factors are important but not directly causal. Genetic influences on delusional ideation have received little specific study. When links between genetic factors and delusional beliefs have been made, these usually emerge as post hoc findings from research protocols designed to study paranoid schizophrenia.

Dopamine Dopamine is no longer viewed as a likely contender for the status of single schizophrenogenic factor. However, dopamine-releasing illicit drugs are thought to trigger delusions, especially delusions of persecution.

There is some evidence of dopaminergic abnormalities specific to delusional ideation. However, biological researchers seem to have unwittingly hampered their own investigations in a potentially important area by studying broad constructs such as paranoid schizophrenia. It may be the case, for instance, that disturbance of mental health in general is too broad a concept to be related to specific neurotransmitter mechanisms even if individual psychotic phenomena, such as hallucinations and delusions, may be. Dopamine functioning remains a possible avenue for the investigation of paranoia.

Psychological Approaches

In 1983, Winters and Neale distinguished between motivational and deficit psychological theories of delusional thinking. Deficit theories generally assume that cognitive or attentional abnormalities lead to delusions, whereas motivational theories, in general, assume that individuals develop delusions to explain either unusual perceptual experiences or aversive psychological states.

Psychological deficit theories

Deficits (or excesses) in emotional expression or processing A number of theorists have suggested that delusions stem from excessive or exaggerated emotion. There is some evidence of disruption (or facilitation) of emotional processing in delusions, particularly paranoia. Generally, research examining cognitive processes in persecutory delusions has examined the processing of information related to emotions and the self-concept.

Deficits in logical thinking Because delusions are illogical, many authors have suggested that delusional beliefs occur as a result of deficits in an individual's ability to think logically. In the presence of schizophrenic thought disorder, normal associative links may be disrupted. In the absence of coherent logical associations, affective influences may become dominant and hence delusional ideas may be generated. Although originally advocated by influential authors, there remain reservations about this very general model.

There have been a number of more specific models of delusions involving deficits in logical processes. In delusions, errors or biases in inductive reasoning have been investigated by a number of researchers. These studies have often been cited as evidence of rapid, overconfident, less evidence-based patterns of inductive reasoning in deluded patients.

Perceptual deficits and excesses It has been suggested that paranoid delusions are the products of healthy rational attempts to explain anomalous or disturbed perceptions. There is some evidence in favor of this model, although some investigations have been somewhat unconventional. It has also been suggested that hallucinations may lead to rational delusional beliefs. It might seem quite rational to develop delusional explanations for inexplicable perceptual experiences. There is some clinical evidence of a relationship between deafness and delusions of persecution in elderly patients, for whom the misperception of the conversations of others may serve as fuel for paranoia. Despite this, and the logic of this

approach, there is little experimental support for such a perceptual deficit model in younger patients. One group of perceptual excesses known to lead to delusions is drug-induced hallucinations. However, the delusions seen in drug use seem qualitatively different from functional psychoses.

Theory-of-mind deficits In 1994, Frith suggested that a deficit or abnormality in what is termed the Theory of Mind may explain a number of schizophrenic symptoms. We do not, as a rule, appraise other people's actions and conversation as simple behavioral or verbal patterns. Rather, according to the Theory of Mind, we assume that we can use behavior to divine wishes, hopes, beliefs, intentions, and so forth. Such an understanding is believed to be vital for successful social intercourse. In Frith's model, paranoid delusions are seen as stemming from a lack of understanding of the actions of other people. Although support for this model comes mainly from phenomenological discussions, a number of recent empirical studies have been reported to support it.

Attentional deficits In 1994, Hemsley suggested that delusions stem from a process involving abnormalities in selective attention in which attention is focused on trivialities. This leads to a lack of influence from learned regularities. Clinically, abnormalities in selective attention seem common. However, certain details of the causal. psychological, linking processes seem vague and poorly related to clinical details of delusional beliefs.

Psychological motivational theories

Learning theory The development of learning theory as a major psychological discipline has, of course, led to behavioral theories of delusions. One general approach is to assume that psychotic phenomena are maintained by positive reinforcement. Indeed, behavioral approaches to the treatment of delusions have frequently proved successful, although this tells us little about the origins of delusional beliefs.

Another behavioral approach to delusions is to assume that delusions are reinforcing through avoidance. Patients with schizophrenia may well be sensitive to anxiety and stress, and therefore that avoidance through delusional ideation may be welcome. Overall, however, these models are unsatisfying. The fact that psychopathology gets worse in conditions of stress is not a good test of any model. Such models seem to invoke, but not address, cognitive variables.

Psychoanalytical theories

The paranoid pseudocommunity Some psychoanalytic theorists have suggested that delusions result from weaknesses in an individual's ego boundaries. If a person's distinction between self and world and between self and fantasy is weak, the distinction between reality and falsehood is assumed to be similarly weak. In this context. it has been suggested that a person frustrated by a lack of social skills seeks refuge in fantasy and daydreams. However, these primitive fantasies, born from frustration, also threaten the ego. Because the defenses (fantasy) are immature, these threats are denied and repressed. Because the individual is not in contact with a normal social network, paranoia develops in a context of imagined and delusional others – the paranoid pseudocommunity.

Although this model appears overly complex, with many assumptions, many researchers have found it persuasive. Empirical data, however, undermine the central thesis that paranoia is born from a lack of social skills. In fact, people diagnosed as having schizophrenia have a higher premorbid social adjustment if they have paranoid symptoms.

Freud More commonly, psychoanalytic theories have invoked defense mechanisms, particularly projection. Freud (in 1911) and many others suggested that delusions represent the externalization of desires, fears, or conflicts. Sigmund Freud's theory of the causes of paranoia has been extremely influential. In essence, it suggests that paranoia develops as a defensive support for a denial of homosexual love. Broad acceptance of this model seems to have shaped American psychiatry's approach to paranoia.

Freud proposed a psychoanalytic theory of persecutory delusions. Paranoid delusions, Freud claimed, are parts of a process protecting the conscious ego from awareness of conflict with unacceptable homosexual impulses stemming from the id. Such homosexual urges are, Freud argued, denied or contradicted and then countered by the defenses of rationalization and projection. A male patient is essentially confronted with the idea that "I (a man) love him (a man)." This is unacceptable, leading to reaction formation and the idea "I do not love him. I hate him." Such inhuman hatred is still unacceptable. and is rationalized as "I hate him because he hates and persecutes me."

Freud's account is colorful, but has been widely challenged. Freud's model suggests that the delusional persecutor is a person for whom the patient expresses secret homosexual love. However, paranoid people frequently do not know their persecutors or even have a specific idea of who they might be. Moreover, many patients fear persecution by members of the opposite sex, who are unlikely to be fueled by repressed homosexual feelings.

It is notable that Freud's model of paranoia includes two elements: a defensive projection or

externalization of threatening material and latent homosexuality. It is possible that the defensive component is more valid than the homosexual part. Psychoanalytic writers after Freud suggested that persecutory delusions serve a defensive function, without necessarily stressing latent homosexuality.

In an more recent variant of this model, Colby suggested that paranoia stems from a tendency of a person to perceive or generate threats to his or her self-esteem combined with a protective mechanism of projection and externalization of the threat to others. In an interesting twist, Colby described a computer simulation model of his theory. A test was conducted comparing the output of his computer model with a real paranoid patient. Because clinical judges were reported to be unable to distinguish the two, the computer model was considered a success. One of the main benefits of Colby's theory is that it is testable. People with delusions of persecution should readily perceive potential threats to self-esteem, and they should also locate the source of such threats as external. Many more recent cognitive investigations of paranoia are compatible with Colby's model.

Camouflaged depression Zigler and Glick suggested that paranoia is a form of camouflaged depression. This suggestion was originally formulated as a statement regarding classification, with paranoia being properly seen as a form of affective disorder. However, the claim is also a causal one. Taking as sources the psychological theories of paranoia previously briefly reviewed, Zigler and Glick claimed that paranoia consists of low self-esteem or feelings of inadequacy (the depression part) covered with defensive projections of these feelings onto others.

Two types of paranoia A subdivision of paranoia has recently been proposed into the rather pejoratively labeled poor-me and bad-me paranoias. In the first form, paranoia is seen as stemming from a low self-esteem, so low that persecution is expected. In the second, paranoia is seen as stemming from defensive processes similar to those suggested by Kinderman and Bentan. Despite weak evidence supporting the subdivision, this categorization has the benefit of being based on theoretical distinctions between different cognitive processes believed to be etiologically important.

Attributional defenses against threats to the self-concept In the attributional model of paranoia, proposed by Kinderman and Bentan, events that might activate an implicit negative self-representations are hypothesized to evoke externalizing (self-protective yet paranoid) causal attributions: "I am not responsible for the bad things that are happening to me, other people are." This reduces the accessibility of self-actual: self-ideal discrepancies, but it leads to the belief that others have negative views of the self. Such an explanation is consistent with psychodynamic accounts of paranoia and with the suggestion that paranoia represents a form of camouflaged depression. This model may have relevance to "normal" psychology beyond the clinical arena.

Normal emotional processing More recently, psychologists Garety and Freeman have suggested that paranoid thought emerges from normal, rather than excessive or exaggerated, emotional processes. These commentators point out that relatively normal emotional problems (anxiety and depression) are common in paranoia and contend that it is possible that these processes may lead to delusional beliefs.

Although it is an attractive argument, this approach fails to account for why some people with these problems remain simply anxious and depressed, whereas other people become deluded.

Treatment of Paranoia

A range of drug treatments are commonly used in the treatment of psychosis, including paranoia. The most important are antipsychotic drugs (sometimes called neuroleptics), which are now regarded as the preferred medical treatment for psychotic conditions. Antipsychotic drugs may be used in the acute phase of psychosis or prophylactically and have been shown to result in acute symptom management and the prevention of relapse. However, neuroleptics are not perfect treatments. Some patients fail to respond or show only a partial response, and neuroleptic treatment can lead to serious and distressing side effects.

Modern developments in clinical psychology mean that schizophrenia may now be seen as amenable to psychological treatment. In particular, delusional ideas can be addressed successfully using the techniques of reality testing and cognitive restructuring. Relatively straightforward cognitive therapy strategies may be beneficial to deluded patients. A common feature of the cognitive-behavioral therapy (CBT) strategies that have been advocated for use with deluded patients is the avoidance of direct confrontation of the patient.

Kinderman and Bentall described one final therapeutic approach, directly driven by theory. In this approach, the causal attributions made by a severely deluded, paranoid patient are addressed through the use of cognitive therapeutic techniques. External personal attributions for minor negative events (blaming other people for bad events) are thought to be characteristic of paranoia. The use of the techniques of CBT may address this tendency.

It has been suggested, with respect to the teleological and religious quality of many delusions, that a

patient's recovery is heralded by the acceptance of the accidental in life. If the situational-external locus of causal attribution is meaningful, it is essentially an acceptance of the accidental. Such an acceptance may be important in many areas of life. Perhaps paranoid individuals need to learn that, sometimes. strange things just happen. An acceptance of the accidental in life may be psychologically important.

See Also the Following Article

Schizophrenia.

Further Reading

Bentall, R. P., Corcoran, R., Howard, R., et al. (2001). Persecutory delusions: a review and theoretical integration. *Clinical Psychology Review* **21**, 1143–1192.

Cole, J., Kleberman, G. L. and Goldberg, S. C. (1964). Phenothiazine treatment of acute schizophrenia. *Archives of General Psychiatry* **10**, 246–261.

Freud, S. (1953–1974). Psycho-analytic notes upon an autobiographical account of a case of paranoia (dementia paranoides) (1911). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 12), London: Hogarth Press and the Institute of Psychoanalysis.

Frith, C. (1994). Theory of mind in schizophrenia. In: David, A. S. & Cutting, J. C. (eds.) *The neuropsychology of schizophrenia*, pp. 147–161. Hove, UK: Lawrence Erlbaum.

Garety, P. A. and Freeman, D. (1999). Cognitive approaches to delusions: a critical review of theories and evidence. *British Journal of Clinical Psychology* **38**, 113–154.

Garety, P. A., Hemsley, D. R. and Wessely, S. (1991). Reasoning in deluded schizophrenic and paranoid patients. *Journal of Nervous and Mental Disease* **179**, 194–201.

Kinderman, P. and Bentall, R. P. (1999). The clinical implications of a psychological model of paranoia. In: Sanavio, E. (ed.) *Behavior and cognitive therapy today: essays in honour of Hans J. Eysenck*, pp. 131–162. Oxford: Elsevier Press.

Winters, K. C. and Neale, J. M. (1983). Delusions and delusional thinking: a review of the literature. *Clinical Psychology Review* **3**, 227–253.

Paraventricular Nucleus

L W Swanson
Neuroscience Research Institute, University of Southern California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by L W Swanson, volume 3, pp 130–133, © 2000, Elsevier Inc.

Location and Basic Divisions
Output Control
CRH Neurons
Stressor-Specific Patterns of Gene Expression
Regulation of Homeostasis and Ingestive Behaviors

Glossary

Corticotrope	Cell type in the anterior pituitary that synthesizes and secretes adrenocorticotropic hormone (ACTH), which in turn stimulates glucocorticoid secretion from the adrenal cortex.
Corticotropin-releasing hormone (CRH)	Peptide hormone that increases the secretion of adrenocorticotropic hormone (ACTH) from corticotropes in the anterior pituitary.
Magnocellular neurosecretory system	Large neurons in and around the hypothalamus that send their axons directly to the posterior pituitary, where they release the peptide hormones vasopressin and oxytocin into the general circulation.
Median eminence	Functionally, the initial part of the stalk of the pituitary gland along the base of the hypothalamus; its inner layer contains the axons of magnocellular neurosecretory neurons and its outer layer contains the axon terminals of parvicellular neurosecretory neurons.
Paracrine	Chemical messengers that act locally in a general way; in contrast to hormonal messengers that act via the bloodstream and to synaptic messengers that act as specialized intracellular junctions.
Parvicellular neurosecretory system	Small neurons in and around the hypothalamus that send their axons to the outer layer of the median eminence, where they are taken up and delivered to the anterior pituitary by a portal system.
Portal system	A system of veins with a capillary network at either end; used to deliver substances from one capillary bed (as in the median eminence) to another (as in the anterior pituitary).

The paraventricular nucleus of the hypothalamus (PVH) is the most important part of the brain for controlling the stress response because its neurons control the release from the anterior pituitary gland of adrenocorticotropic hormone (ACTH), which in turn controls glucocorticoid release from the adrenal cortex. In addition to constituting the final common pathway for all centrally mediated stress responses, the PVH (1) controls the release of other hormones from the anterior and posterior parts of the pituitary gland, (2) influences the size of autonomic reflexes mediated by circuits in the brain stem and spinal cord, and (3) modulates eating and drinking behaviors. Furthermore, steroid hormone and neural inputs can dramatically alter the expression of neuropeptide and many other genes in the PVH, with different stressors producing different patterns of altered expression.

Location and Basic Divisions

In all mammals, the PVH lies adjacent to the third ventricle, in the rostral half of the hypothalamus, rostral to the infundibulum or stalk of the pituitary gland. The PVH, together with the supraoptic nucleus, was originally known as the origin of the magnocellular neurosecretory system. That is, under most conditions separate groups of neurosecretory neurons in each of these two nuclei synthesize significant amounts of either vasopressin or oxytocin and release them into the general circulation at the level of the posterior pituitary gland. These peptide hormones (which are nine amino acids long) play especially important roles in controlling diuresis and blood pressure (vasopressin), and in parturition, lactation, and associated social behaviors (oxytocin).

In the 1970s, it became clear that the PVH also has two other divisions. First, it contains neurons that project to the median eminence and that are thus involved in controlling hormone release from the anterior pituitary (the parvicellular neurosecretory system) and, second, it contains neurons that send descending projections to the brain stem and spinal cord – especially to regions involved in autonomic reflexes, the relay of nociceptive information, and somatic motor behavioral responses. Neuroanatomical studies with multiple retrograde tracers demonstrate that there are three essentially separate cell groups in the PVH of the rat: the magnocellular neurosecretory, parvicellular neurosecretory, and descending divisions.

The parvicellular neurosecretory division appears to contain three major cell types. Most importantly regarding stress, it contains virtually all of the neurons that synthesize corticotropin-releasing hormone (CRH) and release it into the pituitary portal system to increase ACTH secretion from the anterior pituitary. Second, the PVH contains virtually all of the neurons that synthesize thyrotropin-releasing hormone (TRH) and release it into the pituitary portal system to increase thyroid-stimulating hormone secretion from the anterior pituitary. And third, the periventricular zone of the PVH contains a large fraction of the neurons that release somatostatin into the pituitary portal system to inhibit growth hormone secretion from the anterior pituitary.

Neurons in the descending division of the PVH innervate many parts of the brain stem and spinal cord. Best understood, however, are inputs to preganglionic neurons of both the parasympathetic and sympathetic parts of the autonomic nervous system to visceral sensory regions including the nucleus of the solitary tract and parabrachial nucleus, to the periaqueductal gray, and to various parts of the reticular formation. As discussed in the following, these descending projections appear to coordinate autonomic and behavioral responses with neuroendocrine responses elicited from the PVH.

Output Control

As the final common pathway for all types of stress response mediated by the central nervous system, it stands to reason that the PVH receives pathways for all of the relevant stimuli. Thus, it should come as no surprise that the afferent control of the PVH is exceedingly complex. In fact, on the order of 50 different sources of axonal inputs to the PVH have been identified with modern neuroanatomical pathway tracing methods, and in many instances associated neurotransmitters have been identified as well. Each neural input appears to end in relation to more than one cell type in the PVH, and a majority of the neural inputs that have been identified to date end in the region of CRH neurons that control ACTH release.

Neural inputs to parvicellular neuroendocrine CRH neurons fall into a number of functional categories. First, there are inputs that relay sensory information. The best-known inputs of this type arise in three hindbrain regions – the nucleus of the solitary tract, ventrolateral medulla, and parabrachial nucleus – and are known to relay visceroceptive, and probably nociceptive, information to CRH neurons. Second, there is a direct input from a part of the basal ganglia (of the cerebral hemispheres) known as the bed nuclei of the stria terminalis. The bed nuclei in turn receive inputs from other parts of the cerebral hemispheres – including the ventral hippocampus, amygdala, and prefrontal cortex – that may exert cognitive influences on the PVH. Third, there are direct inputs from certain behavioral state control

nuclei, such as the suprachiasmatic nucleus (which generates the circadian rhythm) and the midbrain raphé nuclei (which may be involved in certain aspects of behavioral state or arousal). And fourth, parvicellular neurosecretory CRH neurons in the PVH receive inputs from many cell groups in the hypothalamus that have complex, as yet poorly understood functions related to homeostasis and behavior.

In addition to neural inputs, the PVH appears to be influenced by humoral inputs, that is, by feedback influences of steroid and other hormones. For example, specific sets of neurons in the PVH express glucocorticoid, mineralocorticoid, estrogen, and thyroid hormone receptors. These hormone receptors may change the plasma membrane electrical properties of certain PVH neurons, but it is becoming increasingly clear that they may also change – either directly or indirectly – the pattern of gene expression in PVH neurons in profound ways.

CRH Neurons

CRH neurons that control ACTH secretion are quite remarkable from a functional point of view. In the first place, they have the traditional role of releasing a factor or hormone (CRH) into the proximal capillary plexus of the hypothalamo-hypophysial portal system – at the level of the external layer of the median eminence. The axon of these neurons initially extends laterally from the cell body in the PVH and then arches ventrally through the region of the fornix before curving medially around the ventrolateral border of the hypothalamic ventromedial nucleus to enter the median eminence, where it ends in a rich proliferation of terminal boutons. Peptides (such as CRH) released from these terminals are of course transported via the portal system to act as hormones in the anterior pituitary (for example, on corticotropes).

In addition, however, there is evidence that released substances may also act in a paracrine way within the median eminence itself. For example, it would appear that CRH can act presynaptically to inhibit the release of gonadotropin-releasing hormone (GnRH) at the level of the median eminence. This may account in part for the inhibition of gonadal steroid hormone release observed after some types of stress.

Furthermore, the results of lucifer yellow dye cell-filling experiments indicate that the axon of parvicellular neurosecretory neurons generates terminal boutons of passage in the hypothalamus, on its way to the median eminence. This raises the distinct possibility that these neurons also form classical chemical synapses within the hypothalamus itself. Thus, the evidence suggests that the axon of CRH parvicellular neurosecretory neurons in the PVH can have spatially segregated hormonal, paracrine, and synaptic effects.

In addition to the fact that ACTH-controlling CRH neurons receive more than 50 sources of neural inputs, and that their axon can exert hormonal, paracrine, and synaptic effects, it is now clear that they can also express a bewildering complement of neurotransmitter and/or neuromodulator genes – on the order of 10 to 20. In addition to CRH, they include vasopressin and angiotensin, all of which act synergistically to secrete ACTH, as well as enkephalin, neurotensin, and cholecystokinin, which do not appear to act on corticotropes but may instead interact with receptors at the level of the median eminence and/or hypothalamus.

Stressor-Specific Patterns of Gene Expression

Shortly after CRH was characterized and antibodies to it were raised by Vale and his colleagues, it became clear from multiple labeling immunohistochemical studies that the ratio of putative neurotransmitters in individual PVH neurons can vary dramatically, depending on the endocrine status and stress history of the animal. The was first observed in adrenalectomized rats, in which high levels of previously undetected vasopressin were found in CRH neurons, indicating that vasopressin mRNA levels are much more sensitive to the negative feedback effects of glucocorticoids than CRH mRNA levels.

Since then, a vast amount of experimental evidence has suggested that many forms of stress alter patterns of neuropeptide gene expression in the PVH, as well as in numerous other parts of the brain, and that each form of stress may be accompanied by a specific and distinct pattern of altered gene expression in the brain. For any particular gene, the occurrence, magnitude, and time course of altered expression depend on the type, magnitude, and time course of a particular stressor or pattern of stressors. Thus, it is clear that stressors do not produce synchronous changes in brain gene expression.

Interestingly, a number of neurotransmitter or neurotransmitter-related genes are expressed across cell types in the PVH. For example, under appropriate conditions, CRH, vasopressin, angiotensin, dynorphin, and tyrosine hydroxylase (the enzyme that synthesizes dopamine) are expressed in parvicellular neurosecretory, magnocellular neurosecretory, and descending neurons of the PVH, and a particular stressor or condition may influence gene expression differentially in different cell types. In the rat, corticosterone (the principal endogenous glucocorticoid) decreases CRH and vasopressin mRNAs in parvicellular neurosecretory

neurons, increases CRH mRNA but has no effect on vasopressin mRNA in magnocellular neurosecretory neurons, and increases CRH mRNA but has no effect on vasopressin mRNA in descending neurons.

Regulation of Homeostasis and Ingestive Behaviors

A wide range of experimental work has shown that the PVH is part of the circuit or system that controls eating and drinking behaviors, that is, ingestive behaviors, and that descending projections from the PVH (probably to the periaqueductal gray or nearby areas) are involved in these responses. This involvement in behaviors that play a critical role in maintaining homeostasis for metabolic fuels and for body water is perhaps not surprising in view of the neuroendocrine and autonomic projections of the nucleus. Regarding metabolic control, recall that virtually all parvicellular neurosecretory CRH and TRH neurons (controlling adrenal glucocorticoid and thyroid hormone secretion, respectively) are found in the PVH, as are a majority of the somatostatin neurons that inhibit growth hormone release. And regarding body water regulation, recall that the PVH contains a major population of magnocellular neurosecretory vasopressin (antidiuretic hormone) neurons. In addition, the PVH establishes major bidirectional connections with brain stem and spinal cord circuits that mediate both sympathetic and parasympathetic reflexes, some of which are undoubtedly involved in metabolic and body water regulation.

In summary, the PVH consists of three functionally distinct divisions with essentially motor functions: parvicellular neurosecretory, magnocellular neurosecretory, and descending, with inputs to autonomic and somatomotor control systems. As such, the PVH is in an ideal position to integrate neuroendocrine, autonomic, and behavioral responses to particular stressors, which is a hallmark of hypothalamic function in general. Because the PVH contains few if any interneurons, and recurrent axon collaterals have been difficult to identify unequivocally, this functional integration appears to be mediated by neural inputs, virtually all of which innervate more than one functional division, and perhaps by humoral inputs as well. In addition to containing the CRH neuroendocrine motoneurons that control ACTH release, the PVH is critical for the integrated behavioral, autonomic, and neuroendocrine responses that assure a relatively constant supply of nutrients and body water.

See Also the Following Articles

Homeostasis; Hypothalamic-Pituitary-Adrenal; Pituitary Regulation, Role of.

Further Reading

Herman, J. P., Ostrander, M. M., Mueller, N. K. and Figueiredo, H. (2005). Limbic system mechanisms of stress regulation: hypothalamo-pituitary-adrenocortical axis. *Progress in Neuropsychopharmacology and Biological Psychiatry* **29**, 1201–1213.

Leibowitz, S. F. (1992). Neurochemical-neuroendocrine systems in the brain controlling macronutrient intake and metabolism. *Trends in Neuroscience* **15**, 491–497.

Reyes, T. M., Walker, J. R., Decino, C., Hegenesch, J. B. and Sawchenko, P. E. (2003). Categorically distinct acute stressors elicit dissimilar transcriptional profiles in the paraventricular nucleus of the hypothalamus. *Journal of Neuroscience* **23**, 5607–5616.

Swanson, L. W. (1991). Biochemical switching in hypothalamic circuits mediating responses to stress. *Progress in Brain Research* **87**, 181–200.

Swanson, L. W. and Sawchenko, P. E. (1980). Paraventricular nucleus: a site for the integration of neuroendocrine and autonomic mechanisms. *Progress in Neuroendocrinology* **31**, 410–417.

Watts, A. G. (2005). Glucocorticoid regulation of peptide genes in neuroendocrine CRH neurons: a complexity beyond negative feedback. *Neuroendocronology* **26**, 109–130.

Young, L. J., Murphy Young, A. Z. and Hammock, E. A. (2005). Anatomy and neurochemistry of the pair bond. *Journal of Computational Neurology* **493**, 51–57.

Parenting, Stress of

K D Jennings
University of Pittsburgh, Pittsburgh, PA, USA
LJ Dietz
University of Pittsburgh Medical Center, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Why Parenting Can Be Stressful
Types of Parenting Stress
Effects of Stress on Parents
Effects of Parenting Stress on Children
Interventions for Parents and Children

Glossary

Child temperament	An enduring pattern of biological and physiological responses to the environment.
Coping	Psychological processes that allow an individual to regulate negative emotions and behavior to adapt to stressful situations.
Daily hassles	Normal or typical life experiences that increase negative affect and interpersonal distress by virtue of their high demands and low satisfaction.
Major life events	Exceptional but time-limited life circumstances that produce stress in an individual by acutely increasing negative affect, interfering with daily functioning, and creating change in relationships and routines.
Parenting stress	Pattern of adverse psychological and physiological reactions resulting from the demands of being a parent.
Self-efficacy	An individual's perception of his or her competence in completing the tasks that define a particular domain.
Social support	Interpersonal relationships that meet the emotional and instrumental needs of individuals experiencing distress.

Raising children is often a rewarding and positive experience, but it is also a challenging task that is inevitably stressful at times. Parenting requires a high level of effort and emotional involvement that can be overwhelming. It also involves a wide range of skills that must be used in a flexible manner. Parenting involves meeting the basic needs of children, socializing them to culturally appropriate rules of behavior, and promoting their self-esteem and emotional well-being. In the past 25 years, increasing attention has been paid to the stress of raising children and the effects of parenting stress on caregivers and families. Parenting stress has been defined as a pattern of adverse psychological and physiological reactions resulting from attempts to adapt to the demands of being a parent.

Why Parenting Can Be Stressful

Parenting Tasks

On a daily basis, parents must meet children's survival needs by providing them with shelter, food, clothing, and other necessities. They must ensure their physical safety by carefully monitoring their activities in different environments and by ensuring that proper medical attention is received. Parents must also meet their children's emotional needs by providing attention and affection in order to promote a healthy self concept and socioemotional development. Parents are also responsible for socializing children. They teach them how to adapt to their sociocultural environment by monitoring their behavior and reinforcing culturally appropriate responses. Parents also act as the primary models for regulating negative emotions and coping with distress in culturally specific ways. Parents actively help younger children cope with anger and disappointment by emotional scaffolding and redirecting attention. With older children, parents assist in problem solving and anticipating consequences prior to responding to a situation. Parents provide negative consequences to dissuade inappropriate anger responses such as aggression and prolonged displays of negative affect; in contrast, they praise and reward behaviors that are appropriate or polite. Finally, parents are responsible for introducing children to the larger community outside the family by providing opportunities for peer interaction, formal learning in school, and group activities. In all, parents are expected to carry out a wide variety of tasks that enable their children to successfully adapt to the demands of their sociocultural group and ultimately to become contributing members of their group.

Parenting Must Change as Children Develop

As children grow and develop, their needs and reliance on parents continually change. Being a parent is thus not a singular transition into a static role but rather a set of transitions within a dynamic role that is continually evolving to meet the changing needs of children. Furthermore, the role of being a parent

requires synchrony between parents and children, and this synchrony falls largely on parents to achieve, especially when children are younger. Thus, children's developmental stages influence the amount of stress experienced by parents. For example, caring for a newborn is a time of high stress, particularly for the first-born child. In addition to the change in responsibilities and life style that occurs with the first child, newborns demand round-the-clock intensive caretaking. The preschool age has also been identified as stressful because parents must provide constant care and supervision while negotiating limits with their children as they begin to establish their independence. Other developmental stages provide their own set of challenges: entry into school, with exposure to and judgment by the larger community; adolescence, with increased demands for independence and a renegotiation of the parental role; and finally launching, facilitating the children's assumption of adult roles. Each developmental stage demands different characteristics of the parent; thus, parents differ on which stages they find easier or more difficult.

Parenting Must be Adapted to Children's Temperament and Special Needs

In addition to development, parents must adapt their interactions and interventions according to their child's individual temperament. There is not a single style of parenting that is ideal for every child. Instead, for example, parents must adapt their style to be more soothing for children with difficult temperaments who have problems regulating negative affect and more firm for children with temperaments that are high in novelty seeking. Children with more extreme temperaments or temperaments that differ from their parents can be experienced as more stressful. In addition, disabilities, developmental delays, chronic illnesses, or other special needs of children influence the level of stress experienced by parents. Such children typically require more physical care and more supervision; they also require that parents adjust to the loss of their perfect child and learn to cope with the medical and educational systems. Similarly, children with behavioral or emotional problems cause more stress for their parents.

Balancing Parenting with Other Roles

Parenting stress can also result from difficulty in balancing parenting responsibilities with other adult roles, particularly the roles of spouse and worker. For example, the transition to parenthood has been associated repeatedly with decreased marital satisfaction. The overwhelming, new responsibilities that come from the arrival of a newborn, combined with fatigue and the intense demands of caring for an infant, decrease the time and attention available for the marital relationship, which in turn may decrease marital satisfaction. Over time, parents must learn to balance the needs of all in the family, including the needs of their spouse and their own needs. Parents' difficulty in finding time and energy to meet their own needs is a major contributor to parenting stress.

Balancing work with parenting is another source of stress. Parents must make provisions for routine child care and also accommodate nonroutine events such as children's illnesses and school snow days. The rising demographic trend in first-world nations for dual-wage-earner families indicates that a high number of new mothers will continue to work outside of the home after a child is born. Thus, many families will be faced with the difficulties of balancing multiple roles.

Circumstances of Child Rearing

Specific circumstances of childrearing can exacerbate the level of stress experienced by parents. These circumstances include poverty, number of children, amount of social support, marital status, and physical or mental illness in parents. These stressors impede parents' efforts to care for their children. Thus, for example, single parents raising a number of children in poverty are likely to experience more stress than parents raising children in more favorable circumstances.

Types of Parenting Stress

Major Life Events

Early research on parenting stress focused on the consequences of major life events that increased emotional distress in parents and complicated normal parenting tasks. These major life events, such as death of a spouse, loss of employment, or divorce, were viewed as atypical of the everyday experience of parenting and were thought to increase parental distress and dysfunction and to impair effectiveness of meeting children's needs. Likewise, parenting stress was examined in relation to parenting children with a developmental disability, chronic physical illness, or an emotional disorder. This research did indeed indicate a link between major life events and various indices of parent, child, and family functioning. Parents who were coping with out-of-the-ordinary difficult circumstances reported reduced emotional well-being and increased psychopathology. They were also observed to have less optimal parenting skills. In turn, their children showed more psychopathology and lower social competence. However, a

shortcoming of the major life events model is that it focuses on low-frequency events that cluster in high-risk populations with many problems. The life events model fails to capture the minor stress incurred by all parents in meeting the daily demands of childrearing.

Chronic Minor Stressors

Current research on stress in adults focuses on both major life events and the more mundane chronic stress stemming from meeting the demands of daily living. For parents, minor stressors are the daily hassles of meeting children's needs, including monitoring children's behavior, physically caring for children (i.e., feeding and cleaning), and balancing conflicting schedules of work and family life. Thus, the construct of minor stressors, or parenting hassles, refers to commonly occurring events in all families. Crnic and others argue that the chronic nature of these daily hassles of parenting may make them more likely to adversely affect family functioning than are major life events that occur infrequently. Furthermore, they argue that minor stress, or daily hassles, can provide a more universal model of the effects of stress on family functioning.

Perceived Stress

Stress resulting from daily hassles is largely a subjective experience. Some parents can cope with challenging aspects of parenting better than others. Accordingly, all parents do not experience the daily hassles of parenting in the same way. For example, coping with children's misbehavior is so stressful for some parents that changes occur in the parents' mood and level of anxiety; whereas for other parents these are merely transient annoyances of the day. Individual differences in perceptions of parenting stress reflect both biological and environmental influences. Adults bring to parenting different personality styles, different emotional resources, different coping skills, and different childhood experiences of being parented. Indeed, the best estimate of emotional well-being and functioning in parents is the level of well-being and functioning prior to becoming a parent, that is, prior to the birth of the first child. Personality styles most susceptible to experiencing stress are those marked by the prevalence of negative affect (i.e., moodiness and irritability) and trait anxiety (i.e., intolerance of uncertainty and intolerance of arousal). In addition, parents who as children experienced less positive relationships with their own parents and insecure attachments are more likely to perceive their experiences with their own children as less positive.

Critics of the minor stressors, or daily hassles, approach argue that the construct of daily hassles is inherently subjective, whereas major life events can be assessed more objectively. These critics argue that perceptions of minor hassles are confounded with the outcome measures of family functioning; for example, there is an overlap between perceiving parenting as burdensome and perceiving children's behavior as problematic. On the other hand, individual differences clearly exist in parents' ability to cope with the everyday challenges of parenting. And a large literature verifies that perceptions of events influence functioning.

Effects of Stress on Parents

Parental stress from both major life events and from daily hassles has been shown to negatively impact parents' emotional well-being, marital satisfaction, and parenting behavior.

Parental Mood and Emotional Well-Being

Parenting stress is associated with lower emotional well-being in parents. That is, parents who report more daily hassles in parenting and/or more major life events describe themselves as having less life satisfaction and more negative mood and emotional distress. There is some evidence that stress from daily hassles relates more strongly than does stress from major life events, but stress from each of these sources contributes to lower emotional well-being. In addition to the negative impact on their own emotional well-being, parents with high stress also report a less positive outlook on parenting and less satisfaction in the parental role. Furthermore, they tend to experience less pleasure in and enjoyment of their children. Not surprisingly, they report lower feelings of self-efficacy in the parenting role; that is, they feel less competent in carrying out their parental responsibilities and less confident that their efforts will have a positive impact on their children. For some parents, high levels of parenting stress contribute to psychological disorders, such as depression and anxiety. For example, mothers with higher parenting stress from low-birth-weight or medically ill infants are at higher risk for developing postpartum depression.

A bidirectional relationship exists between parenting stress and negative parental mood. Parenting stress can increase negative parental mood, but, conversely, negative mood may also increase perceptions of parenting stress. Parents with negative mood are more likely to attend to their children's negative behaviors, attribute negative intent to ambiguous behaviors, and have a decreased threshold for aversive behaviors.

Marital Satisfaction and Social Support

The stress of parenting can also affect the marital relationship. When parents feel competent in their abilities to care for their children and cope effectively with daily hassles of parenting, the marital relationship can be strengthened and serve as a solid foundation of the family. However, when high levels of parenting stress are experienced, parents often report lower levels of marital satisfaction. One reason for the decrease in marital satisfaction may be that parents who experience high levels of parenting stress have less time and emotional resources to devote to the marital relationship and to meeting their spouse's needs. Another reason may be that high levels of parenting stress increases the frequency of marital conflict and demonstrations of negative affect toward the spouse. Because the primary social support in parenting is typically the spouse, the decrease in marital satisfaction typically decreases the support a parent experiences in the role of parenting. However, others (such as grandparents) may increase the support that they provide so that a parent does not perceive a decrease in the overall support that he or she receive in the parental role.

Again bidirectional relationships occur. Parenting stress can impair marital satisfaction, but conversely, marital satisfaction can mitigate parenting stress. If parents work together, divide responsibilities, and support one another's efforts, the daily hassles of parenting are less stressful.

Parental Interactions with Children

Parenting stress can also have an adverse effect on parents' behavior toward their children. Parents who experience high levels of stress report more negative and coercive interactions with their children; similarly, they report using more authoritarian and power-assertive discipline strategies. They also describe themselves as less involved with their children. These disparate self-reports are suggestive of inconsistent management techniques. Indeed, high levels of stress have been associated with parents' reports of inconsistent disciplinary styles, that is, alternating between setting harsh and rigid limits for their children to being overly lax and permissive. Although much research has established that stressed parents report negatively on their parenting skills, it is less clear that their actual parenting skills are impaired. Research on observed parenting behaviors has been equivocal, although there is some evidence that mothers are more irritable with their children on days when they have experienced more minor stressors. Thus, parents who report higher levels of stress in caring for children may be less likely to initiate pleasurable interactions with the children and may be less able to use proactive disciplinary strategies.

Effects of Parenting Stress on Children

Parenting stress has repeatedly been found to relate to children's social competence and to behavior problems. Indeed, it is thought to be a primary contributor to the development of psychopathology in children. Again, however, much of this research is based on parental reports. Parents who report high stress describe their children as less socially competent and as having more behavior problems, both externalizing and internalizing problems. These parents also describe their children as exhibiting behaviors consistent with insecure attachments. Limited research using teacher reports and children's self-reports provides some support that parents' negative views may be accurate. Teachers rated the children of highly stressed parents as less socially competent, and children reported more adjustment problems when their parents were highly stressed.

How Might Parenting Stress Affect Children's Outcome?

Parenting stress negatively affects parents' interactions with their children, and the negative interactions, in turn, negatively impact child adjustment. In other words, parenting stress may lead to critical and negative parent–child interactions as well as inconsistent discipline; then these changes in parenting may have negative effects on the children's attachment, self-esteem, and behavior. Another way that parenting stress may affect children's outcome is through modeling affect and behavior regulation. Children observe how parents cope with frustrating situations and how they regulate negative emotions and impulsive behaviors. Parents who experience high parenting stress may model poor regulation of negative affect and more reactive behavioral responses to frustration. Finally, children may be directly experiencing some of the same factors that are causing stress in their parents, such as overcrowding.

Bidirectional effects occur in this arena as well. Parents are crucial in shaping children's behavior, but children also shape adults' parenting behavior. The bidirectional nature of parent–child interactions can create coercive relationships. In these relationships, parents' irritability and less-sensitive parenting fuel behavior problems in children, which evoke further punitive and negative parental behaviors. The cyclical interaction of parents' reactions to stressful circumstances and children's aversive behavior can reinforce a maladaptive pattern of interactions that can lead to

chronic stress in the family environment and can place children at higher risk for delinquency.

Interventions for Parents and Children

With increased attention to the potentially harmful effects of stress on parents, the study of how to promote resiliency in parents has gained increasing attention. Social support and active coping skills have emerged as two primary areas of intervention for relieving parental stress. Social support has been related to decreases in self-reported parental stress. Being connected to others, sharing difficult experiences, and receiving help with the day-to-day care of children reduces parenting stress. Teaching parents strategies for coping with the daily hassles of parenting has also been shown to reduce stress in the parental role. Successful coping involves adapting to the daily hassles of parenting by anticipating problems and having a plan for actively addressing them. For some parents, coping may involve monitoring their children's behavior and environment in order to circumvent the escalation of problem behaviors; for others, it may involve establishing predictable daily routines for children and scheduling periods for taking time off, when another caregiver (e.g., spouse, grandparent, or babysitter) can provide respite care. Specific programs that teach effective parenting skills seem to reduce stress by increasing parents' confidence and efficacy in managing children's problem behaviors. Likewise, interventions that teach young children social problem-solving skills can also decrease parents' stress and increase positive parent–child exchanges. Finally, interventions to decrease parenting stress at the larger community level are also needed to support families with children. Systematic efforts to provide high-quality, affordable child care for working parents, job flexibility for balancing the demands of full-time employment and parenting, and community recreation programs for families will help reduce the burden of parenting and enhance parents' ability to care for their children.

Further Reading

Anthony, L. G., Anthony, B. J., Glanville, D. N., et al. (2005). The relationships between parenting stress, parenting behavior and preschoolers' social competence and behavior problems in the classroom. *Infant and Child Development* **14**, 133–154.

Belsky, J. (1984). The determinants of parenting. *Child Development* **55**, 83–96.

Crnic, K. A. and Acevedo, M. (1995). Everyday stresses and parenting. In: Bornstein, M. H. (ed.) *Handbook of parenting Vol. 4: Applied and practical parenting*, pp. 277–297. Mahwah, NJ: Lawrence Erlbaum.

Crnic, K. A., Gaze, C. and Hoffman, C. (2005). Cumulative parenting stress across the preschool period: relations to maternal parenting and child behavior at age 5. *Infant and Child Development* **14**, 117–132.

Crnic, K. A. and Greenberg, M. T. (1990). Minor parenting stresses with young children. *Child Development* **61**, 1628–1637.

Deater-Deckard, K. (1998). Parenting stress and child adjustment: some old hypotheses and new questions. *Clinical Psychology: Science and Practice* **5**, 314–332.

Deater-Deckard, K. (2004). *Parenting stress*. New Haven, CT: Yale University Press.

Deater-Deckard, K. (2005). Parenting stress and children's development: introduction to the special issue. *Infant and Child Development* **14**, 117–132.

Rodgers, A. Y. (1998). Multiple sources of stress and parenting behavior. *Children and Youth Services Review* **20**, 525–546.

Sepa, A., Frodi, A. and Ludvigsson, J. (2004). Psychosocial correlates of parenting stress, lack of support and lack of confidence/security. *Scandinavian Journal of Psychology* **45**, 169–179.

Parkinson's Disease

S Baser
The Movement Disorder and Spasticity Center, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Clinical Features of Parkinson's Disease
Clues to the Etiology of Parkinson's Disease
Oxidative Stress
Environmental Stresses and Parkinson's Disease
Interactions between Genes and Environment
Medications Used for Parkinson's Disease
Summary

Glossary

Dopamine — A catecholamine neurotransmitter that plays a pivotal role in several key functions of the brain. Normal dopamine neurotransmission in brain is vital for normal motor function. The abnormal movements and deficiency in motor control that characterize Parkinson's Disease are due to a deficiency in central dopamine consequent on the degeneration of dopaminergic neurons.

L-DOPA (Levodopa) — 3,4-dihydroxy-L-Phenylalanine is the immediate precursor of dopamine in its biochemical synthesis. Since L-DOPA can cross the blood brain barrier, it is used clinically in the management of Parkinson's Disease to drive increased dopamine synthesis.

Parkin — An E3 ligase protein that normally tags alpha synuclein protein molecules with ubiquitin. This Parkin-induced tagging (ubiquination) of alpha-synuclein, triggers alpha-synuclein degradation by the cell proteasome. In many cases of inherited Parkinson's disease, however, Parkin is abnormal and fails to ubiquinate alpha synuclein. As a consequence, synuclein does not undergo degradation. Rather, it accumulates in Lewy bodies, which leads to dopaminergic cell dysfunction and death.

Parkinson's disease (PD) — The second most prevalent neurodegenerative disorder of the Western world that leads to abnormal movements and loss of motor control and function. Specifically, the clinical features include slowness of movement (bradykinesia), muscular rigidity, resting tremor and postural instability. PD is characterized by the appearance of intracytoplasmic inclusions called Lewy bodies in dopamine neurons in the substantia nigra and the progressive loss of these neurons. Lewy bodies are enriched in filamentous α-synuclein and other proteins that are often highly ubiquitinated. The majority of PD cases are sporadic and of unknown origin, but several genes have been identified that, when mutated, give rise to rare, familial forms of the disease.

Substantia nigra — A midbrain nucleus of the basal ganglia complex that controls motor function. The pars compacta of the substrantia nigra contains the cell bodies of the dopaminergic neurons that project to the striatum in the forebrain. Degeneration of these neurons is the most prominent neuropathological feature of Parkinson's Disease and is thought to be the prime cause of the disorder.

Synucleins — Small, soluble proteins expressed primarily in neural tissue and in certain tumors. The family includes three known proteins: alpha-synuclein, beta-synuclein, and gamma-synuclein. Mutations in alpha-synuclein are associated with rare familial cases of early-onset Parkinson's disease, and the protein accumulates abnormally in Parkinson's Disease, Alzheimer's Disease, and several other neurodegenerative illnesses. Alpha synuclein has also been implicated in non-familial (sporadic) PD, but the mechanism has yet to be determined.

Ubiquitin proteasome system — Molecular system by which misfolded proteins are tagged with polyubiquitin for subsequent degradation by the proteasome of the cell.

Clinical Features of Parkinson's Disease

Parkinson's disease (PD) is a slowly progressive, degenerative, central nervous system disorder. First described by James Parkinson in 1817, it is a movement disorder that results from the death of dopamine neurons in the substantia nigra. The substantia nigra is located in the midbrain region of the brain. It consists of two parts: the pars compacta (SNc) and the pars reticulata (SNr). The SNc produces dopamine and is the part that degenerates in PD. After 50% of the dopamine neurons and 75–80% of striatal dopamine is lost, patients start to show symptoms

of PD. Associated with the loss of dopamine is the presence of protein-containing deposits known as Lewy bodies. It is not known what actually initiates the development of PD.

Approximately 1 million patients worldwide have PD, with approximately 60 000 new patients diagnosed per year. The average age of onset is 60 years old; 5–10% show symptoms prior to age 40; 3% show symptoms after age 65, and 10% show symptoms after age 80.

To be diagnosed with PD, the patient must have at least three of four cardinal signs: rest tremor, bradykinesia (slowness of movement), rigidity, and loss of balance. Pharmacologically, PD is differentiated from clinically similar disorders by the absence of secondary causes and by the presence of a good response to the medication L-dopa. Definitive diagnosis of PD can only be made by autopsy.

Parkinsonism

Parkinsonism is a clinical syndrome that may share some features with PD, secondary to a variety of etiologies. These etiologies include infectious, vascular, pharmacological, toxic, metabolic, structural, and various degenerative disorders.

Clues Suggesting Atypical Parkinsonism

- Early onset of, or rapidly progressing, dementia
- Rapidly progressive course
- Supranuclear gaze palsy
- Upper motor neuron signs
- Cerebellar signs: dysmetria or ataxia
- Loss of bladder/bowel control
- Early symptomatic low blood pressure
- Early loss of balance

Clues to the Etiology of Parkinson's Disease

For decades scientists knew that the substantia nigra in the midbrain of Parkinson's patients show loss of dopamine nerve cells, but what caused this loss remained elusive. Von Economo's encephalitis, which started as an epidemic in Europe in 1918 and then spread worldwide was associated with symptoms similar to PD in 80% of survivors. This condition was termed postencephalitic parkinsonism.

An important breakthrough in the pathophysiology of PD occurred in the late 1950s when researchers gave rats powerful antipsychotic drugs that blocked dopamine nerve cells. In the 1960s, scientists used the toxin 6-hydroxy-dopamine in animals to destroy these cells and produce a rat model of PD.

Finding effective treatments for PD proved more difficult because dopamine does not pass directly into the brain. As depicted in the movie *Awakenings*, doctors in the late 1960s began to give patients the dopamine precursor L-dopa, which crosses the blood–brain barrier and is converted into dopamine.

In the 1980s, a sudden outbreak of severe PD appeared in young Californian drug addicts. These addicts had used a drug contaminated with 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP). This chemical is converted to MPP^+ in the brain, damaging and killing dopamine cells in the substantia nigra. This inadvertent discovery created a second animal model with which to explore the disease.

Oxidative Stress

Damage from oxygen may be one cause of PD. Oxygen free radicals are oxygen molecules that carry an extra electron. They are destructive because in excessive amounts they chemically attack the components of the cell, including proteins, mitochondria, DNA, and lipids in cell membranes. One of the major problems of normal aging is an increased level of these free-radical-damaged proteins, along with damaged DNA and lipids.

Recent research by Bennett at the University of Virginia isolated complex I from the mitochondria of the brains of PD patients. They discovered that the complex I assembly in Parkinson's had 50% more damage from oxygen. The complex I in the brains of PD patients also had evidence of not being properly assembled and had reduced electron flow.

Environmental Stresses and Parkinson's Disease

Toxins and Parkinson's disease

Rural living and pesticide exposure Although studies in the past pointed to an association between rural living and PD, the results of the most recent studies are quite inconsistent. Because there are many specific exposures associated with living in rural areas, recent studies have attempted to measure these exposures in order to explain the association between PD and areas of low population density.

Three recent studies have found an association between PD and agricultural work. A German study reported an elevated and significant relative risk associated with mushroom harvesting during childhood and adolescence, but no association was found with previous farm activity or employment in agricultural work, living on or near a farm, having contact with farm animals, or being involved in slaughtering

animals. Well water has also been implicated in PD; recent studies in Italy and Spain found a positive association between drinking well water and the risk of PD.

The similarity between the structures of the MPP^+ and the herbicide paraquat encouraged the investigation of a possible relation between pesticide exposure and PD. A study in Taiwan indicated that the incidence of PD was two times greater in people who had used both paraquat and herbicides/pesticides than in those exposed to pesticides and herbicides alone.

Previous studies found a substantially increased rate of PD among city dwellers who gardened as a hobby. Recent studies have consistently shown an increased risk of PD with pesticide exposure, and in some, this achieved statistical significance. A 2000 California-based study showed that individuals with high-level herbicide exposure had a 70% increased risk of developing PD, compared with those who were not exposed. People who used insecticides in the garden showed a 50% increased risk of developing PD compared to those who had never been exposed to home pesticides of any type. In-home use of insect-killing chemicals was associated with a 70% increased risk of developing PD compared with no use of pesticide.

Metal/industrial exposure Several early case reports described Parkinsonian symptoms in association with manganese mining fumes and accidental industrial poisoning. A recent study of 15 welders with Parkinsonian symptoms showed that the welders had an earlier onset of symptoms. Another study reported that counties in Michigan with iron and copper industries had higher PD mortality rates.

Higher mortality rates from PD are reported in regions with paper and chemical industries. More PD patients than controls reported that they had wood paneling in their homes; this may implicate wood preservatives in the etiology of PD.

Head Injury

Head injury has been implicated in the etiology of PD, possibly through mechanisms that are involved in the inflammatory process. Several recent studies found significant risks associated with head trauma, and several well-known boxers are known to have PD.

Smoking

Many epidemiological studies have shown smoking to be protective for PD. A recent prospective study involving the Honolulu Heart Study supported the idea that nicotine may be protective for PD. Other case–control studies, however, do not support the claim that smoking is protective for PD; the previous negative association was explained by the fact that smokers die sooner than nonsmokers.

Diet and Antioxidants

A low-calorie diet may help to reduce the risk of PD. A team from the U.S. National Institute on Aging found a long-term reduction in caloric intake protects rhesus monkeys from developing the disease; they also had higher levels of a growth factor, glial cell line-derived neurotrophic factor (GDNF), which may protect brain cells from destruction.

The role of dietary antioxidants has been examined in the etiology of PD. According to the oxidative stress model, an increase in antioxidants prevents damage and death to the dopaminergic cells by scavenging more free radicals. Epidemiological studies examining the association between antioxidant consumption, including vitamin E, C, and A, have not shown a consistent role.

Vitamin B_6 is essential for the metabolism of protein and for brain function. In a study of more than 5000 people, Dutch researchers found those who reported taking in the most vitamin B_6 were approximately half as likely to develop PD as those who consumed the least of this vitamin.

High levels of the amino acid homocysteine can damage brain cells. To see whether a higher intake of folate and vitamins B_6 and B_{12}, which can reduce homocysteine levels, also reduced the risk of PD, researchers followed 5289 men and women ages 55 and older. The researchers found no association between the consumption of B_{12} or folic acid and the risk of developing the disease; however, the more B_6 people consumed, the lower their risk of PD.

Interactions between Genes and Environment

Genes, Cell Death, and α-Synuclein

No one single underlying factor has been shown to cause PD. Most current research show that it is a complex multifactorial disease resulting from interaction between one or more genes and the environment. The genes that cause inherited PD can involve the proteins α-synuclein, parkin, and ubiquitin. α-Synuclein forms a major component of Lewy bodies, a classic pathological finding in PD patients.

The discovery of the dominantly inherited gene (α-synuclein, SCNA gene, PARK1) was first described in the Contursi region of Italy. This discovery developed the recognition of a direct role for genetics in PD. Abnormally high levels of α-synuclein, which is produced in dopaminergic cells, may play an

Table 1 Medications[a]

Medication	*Mechanism*	*Examples*
L-dopa	Crosses the blood–brain barrier and is converted into dopamine	
MAOIs	Block dopamine metabolism	Selegiline, rasagaline
Dopamine agonists	Act directly on dopamine receptors (D1 and D2) to take the place of dopamine	Bromocriptine (Parlodel), pergolide (Permax), pramipexole (Mirapex), ropinirole (Requip), apomorphine
NMDA antagonists	Block excitatory receptors	Amantadine and flumadine
COMT inhibitors	Block the breakdown of L-dopa in the bloodstream, allowing more L-dopa to cross the blood–brain barrier	Tolcapone (Tasmar), entacapone (Comtan)
Anticholinergics	Decrease the activity of acetylcholine (a balancing neurotransmitter)	Trihexyphenidyl (Artane), benztropine mesylate (Cogentin)

[a]COMT, catechol-O-methyltransferase; MAOIs, Monoamine oxidase inhibitors; NMDA, N-methyl-D-aspartic acid.

important role in the death of these cells. In patients with the defective gene, synuclein accumulates, causing dopamine neurons to become more vulnerable to neurotoxic insults. Normally, two other molecules, parkin and ubiquitin, are involved in the natural self-destruction of synuclein in the natural process of programmed cell death called apoptosis.

Genes, Cell Death, and Ubiquitin

Recessively inherited juvenile parkinsonism (AR-JP, PARKIN, PARK2) is one of the most common forms of familial PD. AR-JP is characterized by the selective and massive loss of dopamine neurons in the substantia nigra of the midbrain and the absence of Lewy bodies (the pathological hallmark of PD). Parkin mutations in this form of PD cause the failure of proteolysis mediated by the ubiquitin system and the accumulation of protein, causing dopamine nerve cell death.

Emotional Stress and Parkinson's Disease

Emotional stress increases the symptoms of PD, including the worsening of tremor. Depression in Parkinson patients is common, with some estimates reaching 50% of those affected. Biochemical changes in the brain involving serotonin, catecholamines, and dopamine and emotional responses to chronic disease play a role.

Encouraging patients to perform as many daily activities as possible and to follow a program of regular exercise can help people with PD maintain mobility. Recent research suggests additional benefits of exercise, including neuroprotective effects. Physical and occupational therapy can help patients maintain or regain muscle tone, maintain range of motion, and learn adaptive strategies. Music therapy has been shown to improve slowness in PD.

Caring for a spouse with PD is associated with emotional and social stress, underlining the importance of also assessing the needs of caregivers. One study showed that spouses of PD patients had more severe depression and that a higher proportion reported tiredness, sadness, and less satisfaction with life than healthy elderly subjects. Caregivers can benefit from learning about the physical and psychological effects of PD and about ways to enable patients to function as well as possible. Because such care is tiring and stressful, caregivers may benefit from support groups.

Medications Used for Parkinson's Disease

The medications used for PD patients compensate for the loss of the dopamine-producing cells by either increasing dopamine levels or simulating the effect of dopamine in the brain (Table 1). Dopamine cannot be given directly because it does not cross the blood–brain barrier.

Summary

Many stressors have been associated with the development of PD. Complex interactions among the environment, oxidants, toxins, genes, and emotions play a role in the etiology of PD.

See Also the Following Articles

Brain Trauma; Diet and Stress, Non-Psychiatric; Dopamine, Central; Environmental Factors; Environmental Stress, Effects on Human Performance; Oxidative Stress; Renal and Adrenocortical Actions of Dopamine; Smoking and Stress; Diet and Stress, Psychiatric; Early Environment and Adult Stress; Neurodegenerative Disorders.

Further Reading

Aarsland, D., Larsen, J. P., Karlsen, K., et al. (1999). Mental symptoms in Parkinson's disease are important contributors to caregiver distress. *International Journal of Geriatric Psychiatry* **14**(10), 866–874.

Fearnley, J. M. and Lees, A. J. (1991). Ageing and Parkinson's disease: substantia nigra regional selectivity. *Brain* **114**(5), 2283–2301.
Gosal, D., Ross, O. A. and Toft, M. (2006). Parkinson's disease: the genetics of a heterogeneous disorder. *European Journal of Neurology* **13**(6), 616–627.
Kontakos, N. and Stokes, J. (2000). Monograph series on aging-related diseases. XII: Parkinson's disease-recent developments and new directions. *Chronic Diseases in Canada* **20**(2), 58–77.
Lang, A. E. and Lozano, A. M. (1998). Parkinson's disease, I. *New England Journal of Medicine* **339**(15), 1044–1053.
Lang, A. E. and Lozano, A. M. (1998). Parkinson's disease, II. *New England Journal of Medicine* **339**(16), 1130–1143.
Marsden, C. D. (1994). Clinical Neuropharmacology. *Journal of Neurol Neurosurg Psychiatry* **57**(6), 672–681.
Olanow, C. W. and Tatton, W. G. (1999). Etiology and pathogenesis of Parkinson's disease. *Annual Review of Neuroscience* **22**, 123–144.
Pacchetti, C., Mancini, F., Aglieri, R., et al. (2000). Active music therapy in Parkinson's disease: an integrative method for motor and emotional rehabilitation. *Psychosomatic Medicine* **62**, 386–393.
Polymeropoulos, M. H. (1998). Autosomal dominant Parkinson's disease and alpha-synuclein. *Annals of Neurology* **44**(3 supplement 1), S63–S4.
Riess, O. and Kruger, R. (1999). Parkinson's disease – a multifactorial neurodegenerative disorder. *Journal of Neural Transmission* **56**(supplement), 113–125.
Tanaka, K., Suzuki, T., Chiba, T., et al. (2001). Parkin is linked to the ubiquitin pathway. *Journal of Molecular Medicine* **79**(9), 482–494.
Trimmer, P. A., Swerdlow, R. H., Parks, J. K., et al. (2000). Abnormal mitochondrial morphology in sporadic Parkinson's and Alzheimer's disease cybrid cell lines. *Experimental Neurology* **162**(1), 37–50.
Wersinger, C. and Sidhu, A. (2006). An inflammatory pathomechanism for Parkinson's disease? *Current Medicinal Chemistry* **13**(5), 591–602.

Peacekeeping

B Litz and S Maguen
VA Boston Healthcare System, Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC), and Boston University School of Medicine, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B Litz and E Bolton, volume 3, pp 134–137, © 2000, Elsevier Inc.

Types of Peacekeeping Missions
Unique Stressors of Peacekeeping
Psychological Impact of Peacekeeping
Conclusions and Recommendations

Glossary

Peace enforcement	A process by which military personnel necessitate a cessation of hostilities and facilitate a peace process.
Peacekeeping	Maintaining a military presence that increases the likelihood of peace and reduces the likelihood of fighting among two or more formerly warring parties.
Peacekeeping roles	Duties such as guarding and patrolling vulnerable areas, helping maintain crowd control, operating checkpoints, providing humanitarian assistance, and facilitating the rebuilding of infrastructure, all of which can expose peacekeepers to potentially dangerous and life-threatening situations.
Posttraumatic stress disorder (PTSD)	A psychological disorder that may result from exposure to severe stress and that involves re-experiencing of the traumatic event, avoidance of cues associated with the event, emotional numbing, and hyperarousal symptoms.
Role conflict in peacekeepers	Conflict resulting from exposure to competing demands of neutrality and restraint in the face of life threat or danger, which may lead to feelings of helplessness or frustration.
Rules of engagement	Rules prohibiting use of weapons except in extreme circumstances, which is quite different from ways in which soldiers are traditionally trained to handle provocation and life threat.

Types of Peacekeeping Missions

Peacekeeping has changed considerably following the inception of the United Nations in 1945. There have been 60 peacekeeping missions since 1948, and currently there are 17 active operations throughout the world. Historically, the role of the military in peacekeeping operations has been that of maintaining a strictly neutral presence by overseeing peace accords between formerly warring parties. However, post-Cold War, peacekeepers have shifted to enforcing peace in the midst of active hostilities.

These modern missions have placed peacekeepers into more dangerous and conflict-laden environments. Modern missions require soldiers to be combat trained, yet able to model restraint and maintain neutrality. As a result, peace enforcement is one of the most stressful duties a soldier can be exposed to in the modern military, and the possibility of exposure to traumatic events and death has become even more likely in modern missions. For example, Seet and Burnham found that more deaths have occurred among UN peacekeepers in the past decade than in the previous 40 years, and that deaths from hostile acts have significantly increased following the Cold War.

This article provides brief summaries of the stressors associated with peacekeeping in three different peacekeeping missions, briefly reviews recent peacekeeping missions, summarizes the unique stressors of peacekeeping, and highlights the psychological impact of these missions.

The Multinational Force and Observers in the Sinai

Since the establishment of the state of Israel in 1948, the UN and other multinational organizations have engaged in an ongoing struggle to create and maintain peace between Israel and its neighbors. An example of a very successful attempt to maintain peace in this region is the Multinational Force and Observers (MFO) in the Sinai. Due to the stability of the peace between Israel and Egypt and the receptiveness of both parties to the presence of the MFO, peacekeepers in the Sinai have been exposed to few potentially traumatizing events directly related to their service as peacekeepers. The MFO is an example of a successful peacekeeping mission in which the duty of peacekeepers is strictly to observe, monitor, and report transgressions.

Operation Restore Hope and Operation Continue Hope in Somalia

As a result of the civil war that erupted in Somalia in 1991, humanitarian relief efforts aimed at curbing famine and the spread of disease were interrupted and sabotaged. The UN, with extensive support from the United States, decided to guarantee the provision of humanitarian aid as well as to enforce the peace in Somalia. During this mission, peacekeepers became peace enforcers, and the risk of exposure to potentially traumatic events grew exponentially. Although Operation Reserve Hope (ORH) was a great success with regard to the provision of medical and food supplies, the mission is often considered to be a failure because Somalia continues to be at risk for the devastating effects of famine, political instability, and violence.

While in Somalia, U.S. military personnel were assigned to a variety of tasks, ranging from policing duty to combat-like duty (e.g., patrols, disarming civilians). Peacekeepers were exposed to a fairly well armed civilian population who were actively engaged in interclan war. Peacekeepers in Somalia were also subjected to acts of aggression by some Somalis, yet the strict rules of engagement sharply restricted options for protection or retaliation, creating a general sense of threat.

Operation Joint Endeavor and Operation Joint Guard

In 1991, following the declaration of independence of Slovenia and Croatia, a civil war erupted in the former Yugoslavia. The disillusion of the state gave rise to power struggles over governing control of the remaining states and to old race-based hatreds. Horrible atrocities were perpetrated, including genocidal acts. In response, approximately 60 000 North Atlantic Treaty Organisation (NATO) and U.S. military personnel were deployed as peacekeepers. Peacekeepers in Bosnia-Herzegovina have been subjected to shelling and sniper attacks, and a minority of peacekeepers have also been taken hostage. Some UN peacekeepers have had to stand by helplessly while atrocities were taking place.

Recently Established Peacekeeping Missions

Locations of recent UN peacekeeping missions include (1) Sudan (UNMIS), which began in March 2005 to support implementation of the Comprehensive Peace Agreement, (2) Burundi (ONUB) since May 2004, (3) Haiti (MINUSTAH) since April 2004, (4) Cote D'Ivoire (UNOCI) since February 2004, (5) Liberia (UNMIL) since September 2003, (6) East Timor (UNMISET) since May 2002, and (7) Ethiopia and Eritrea (UNMEE) since September 2000.

Unique Stressors of Peacekeeping

Peacekeepers are at risk for exposure to stressors that are typical of a war zone (artillery fire, land mines, small arms fire). Even when peacekeepers are tasked with maintaining a firm peace, they are faced with the possibility of life threat. They may be fired upon as a result of a misunderstanding, accidentally in crossfire between two armed feuding parties, or during firing close, which occurs when the opponent wishes to intimidate the peacekeepers in order to keep them away from a certain area. Peacekeepers may also witness atrocities committed against fellow peacekeepers and civilians and the malicious destruction of property. Other types of stressful experiences for peacekeepers are boredom, isolation, family separation, exhaustion, unfavorable climatic conditions, and demoralization about the mission's efficacy.

Role conflicts may also provoke additional stress in peacekeepers. For combat-trained soldiers from larger nations, peacekeeping duty may feel incongruent with their training. For example, peacekeepers, unlike traditional soldiers, are often restrained from taking offensive action in conditions of life threat. In addition, the types of defensive military structures that are commonplace in war are often not as available on peacekeeping missions due to the proximity that is required in order to provide humanitarian assistance and protection. The emphasis on proximity rather than protection creates considerable hypervigilance and arousal in peacekeepers and contributes to a general sense of fear. Peacekeepers who are unclear about how to respond to threats and/or experience repeated threats of injury, with little or no opportunity for recourse, are likely to experience great anxiety. Furthermore, peacekeepers who are forced to suppress their frustration, fear, and anger are at risk for acting out their feelings during a mission and/or upon their return home.

Psychological Impact of Peacekeeping

Although there are many potentially traumatizing experiences and psychological conflicts associated with being a peacekeeper, the great majority of soldiers appear to adapt well. However, a sizeable number of peacekeepers are at risk for the development of psychopathology related to their deployment. The most frequent mental health problems associated with peacekeeping are PTSD, depression, anger and hostility problems, and alcohol abuse. Mehlum and Weisaeth reported that 7 years following service, 16% of prematurely repatriated soldiers from southern Lebanon met criteria for PTSD. Passey found that more than 20% of the soldiers deployed as peacekeepers to Bosnia endorsed symptoms of PTSD and depression. In Litz and colleagues' examination of U.S. military personnel who deployed to Somalia, 25% of Somalia veterans reported clinically significant psychological distress, particularly hostility and anger problems, and 8% reported clinically significant PTSD; Dirkzwager and colleagues found similar rates of PTSD among peacekeepers who served in Yugoslavia. Among Kosovo peacekeepers, Maguen and colleagues found that 4% endorsed clinically significant PTSD symptoms; Dirkzwager and colleagues found similar rates of PTSD among peacekeepers who served in Cambodia. Gray and colleagues found that although most follow a standard pattern of PTSD development, some peacekeepers exhibit delayed-onset PTSD after a period of minimal distress.

Adler and colleagues examined factors that predicted PTSD in U.S. military personnel deployed to the Bosnia peacekeeping mission and found that longer deployments and first-time deployments were associated with an increase in distress scores. Interestingly, the association between deployment length and distress was significant only for male soldiers. In a prospective study of Kosovo peacekeepers, Maguen and colleagues found that pre-existing stress symptoms and exposure to stressful war zone events (e.g., going on dangerous patrols) were the most robust predictors of PTSD symptoms. Similarly, in Litz and colleagues' study of Somalia peacekeepers, the extent of exposure to stressful war zone events and frustrations with aspects of the peace enforcement mission (e.g., restrictive rules of engagement) predicted the severity of PTSD. The relationship between war zone exposure and PTSD was strongest for Somalia veterans who had high levels of frustration with the negative aspects of peacekeeping duty. Additional variables associated with PTSD symptoms reported by Dirkzwager and colleagues were feelings of powerlessness and threat, belief that the mission had become meaningless, and lack of control during deployment. Litz and colleagues also found that cohesion and morale during deployment were negatively associated with symptoms of psychological distress and PTSD. On the other hand, Michel and colleagues found that postdeployment adversities most strongly contributed to poor mental health among Swedish peacekeepers that deployed to Bosnia.

Peacekeeping duty can also promote positive outcomes. Dirkzwager and colleagues found that the majority of peacekeepers reported positive consequences of their deployment, with 82% reporting a broadening of their horizon and 52% reporting increased self-confidence. Britt and Adler, and Bartone documented that following deployment to Bosnia, many peacekeepers reported an increase in political understanding, stress tolerance, and professional qualifications. They also found that soldiers who identified more closely with the role of peacekeeper and who believed their assignment to be important were more likely to report a perceived benefit from their peacekeeping experience and were less likely to report burnout and adverse psychological consequences. Finally, the researchers noted that soldiers who reported benefits as a result of the deployment had at least some exposure to the physical damage caused to the local civilians and soldiers from other nations.

Conclusions and Recommendations

Even in the context of conflicts that arise as a result of counterterrorism, peacekeeping and peace enforcement missions will thankfully outnumber wars internationally. Although the psychological casualties

resulting from peacekeeping stress do not compare to combat, it is important to appreciate the mental health needs of soldiers who return from peacekeeping missions, particularly ones that entail unforeseen escalations in hostilities. Peacekeepers may experience a wide range of potentially traumatic events and numerous uniquely stressful experiences. Peacekeeping, especially peace enforcement, missions should include provisions for secondary prevention interventions for those soldiers most at risk for chronic mental health problems following homecoming.

See Also the Following Article

War-Related Posttraumatic Stress Disorder, Treatment of.

Further Reading

Adler, A. B., Huffman, A. H., Bliese, P. D. and Castro, C. A. (2005). The impact of deployment length and experience on the well-being of male and female soldiers. *Journal of Occupational Health Psychology* **10**, 121–137.

Britt, T. W., Adler, A. B. and Bartone, P. T. (2001). Deriving benefits from stressful events: the role of engagement in meaningful work and hardiness. *Journal of Occupational Health Psychology* **6**, 53–63.

Dirkzwager, A. J., Bramsen, I. and van der Ploeg, H. M. (2005). Factors associated with posttraumatic stress among peacekeeping soldiers. *Anxiety, Stress, and Coping* **18**, 37–51.

Gray, M. J., Bolton, E. E. and Litz, B. T. (2004). A longitudinal analysis of PTSD symptom course: delayed-onset PTSD in Somalia peacekeepers. *Journal of Consulting and Clinical Psychology* **72**, 909–913.

Litz, B. T., King, L. A., King, D. W., Orsillo, S. M. and Friedman, M. J. (1997). Warriors as peacekeepers: features of the somalia experience and PTSD. *Journal of Consulting and Clinical Psychology* **65**(6), 1001–1010.

Litz, B. T., Orsillo, S. M., Friedman, M., Ehlich, P. and Batres, A. (1997). Post-traumatic stress disorder associated with peacekeeping duty in Somalia for U.S. military personnel. *American Journal of Psychiatry* **154**, 178–184.

Maguen, S. and Litz, B. T. (2006). Predictors of morale in U.S. peacekeepers. *Journal of Applied Social Psychology.* **36**, 820–836.

Maguen, S., Litz, B. T., Wang, J. L. and Cook, M. (2004). The stressors and demands of peacekeeping in Kosovo: predictors of mental health response. *Military Medicine* **169**, 198–206.

Mehlum, L. and Weisaeth, L. (2002). Predictors of posttraumatic stress reactions in Norwegian U.N. peacekeepers seven years after service. *Journal of Traumatic Stress* **15**, 17–26.

Michel, P. O., Lundin, T. and Larsson, G. (2003). Stress reactions among Swedish peacekeeping soldiers serving in Bosnia: a longitudinal study. *Journal of Traumatic Stress* **16**, 589–593.

Passey, G. and Crocket, D. (1995). *Psychological consequences of Canadian UN peacekeeping in Croatia and Bosnia*. Paper presented at the meeting of the International Society for Traumatic Stress Studies, Boston, MA.

Seet, B. and Burnham, G. M. (2000). Fatality trends in United Nations Peacekeeping Operations, 1948–1998. *Journal of the American Medical Association* **284**, 598–603.

Peptides

E P Zorrilla and G F Koob
The Scripps Research Institute, La Jolla, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E P Zorrilla and G F Koob, volume 3, pp 138–141, © 2000, Elsevier Inc.

Definition and Classification of Peptides
Biosynthesis and Processing of Peptides: The Proopiomelanocortin Model
Peptide Release
Peptide Receptors
Modulators of Peptide Action

Glossary

Hormone	A peptide, polypeptide, or protein that is released into the circulation via endocrine or exocrine routes to signal to distant tissues.
Neurotransmitter	A compound synthesized and stored within a neuron that, on physiological release into the synapse via a calcium-dependent mechanism, elicits a postsynaptic, receptor-mediated response. Neurotransmitters are also subject to degradation.
Prepro-hormone	A multipotential precursor protein that contains the amino acid sequence for several different peptides that can be

	liberated with appropriate enzymatic processing; typically, it is inert.
Receptor desensitization	The process by which the repeated stimulation of a receptor with an agonist elicits a smaller response or requires larger doses to achieve a similar response.
Ribosome	The structure outside the nucleus of the cell in which mature mRNA is translated into the coded protein.

Definition and Classification of Peptides

Peptides are generally defined as any member of a class of low-molecular-weight compounds that yield two or more amino acids on hydrolysis. Conversely, peptides also can be defined as any compound produced by amide formation between a carboxyl group (COOH) of one amino acid and an amino group (NH_2) of another, thereby resulting in the loss of water from the adjacent amino acids. Although the word peptide is most often used to refer to compounds in which the amide bonds are formed between C-1 of one amino acid and N-2 of another (or eupeptide bonds), it also applies to compounds with residues linked by other amide bonds (or isopeptide bonds). Because the elements of water are removed from amino acids in forming peptide bonds, what remains of each amino acid is referred to as an amino acid residue (e.g., glutamic residue instead of glutamic acid). The 21 amino acids known to be encoded genomically and their common abbreviations are listed in Table 1.

Peptides act as signaling agents to transfer information between tissues. Naturally occurring peptides consist of 100 or fewer residues (longer sequences generally are referred to as proteins and have a more complex tertiary structure), making them up to 50 times larger than classic neurotransmitters. Although peptides are found in much lower concentrations than conventional neurotransmitters, their larger size offers a greater capacity for encoding receptor-appropriate structural information, such that peptides can bind with 1000-fold higher affinity and greater selectivity than their smaller transmitter cousins. Peptides generally colocalize with other peptides and transmitting agents, such as classical neurotransmitters and cytokines. The diversity of available chemical messengers greatly increases the permutations of information transfer from a given cell. Peptides are thought to represent an evolutionarily older class of transmitter than the biogenic amines, with peptides having their roots in hormonal function. More than 70 peptides

Table 1 Amino acids and their symbols

Amino acid	*One-letter symbol*	*Three-letter symbol*	*DNA codons*
Alanine	A	Ala	GCT, GCC, GCA, GCG
Aspartic acid or asparagine[b]	B	Asx	See specific entries
Cysteine	C	Cys	TGT, TGC
Aspartic acid	D	Asp	GAT, GAC
Glutamic acid	E	Glu	GAA, GAG
Phenylalanine	F	Phe	TTT, TTC
Glycine	G	Gly	GGT, GGC, GGA, GGG
Histidine	H	His	CAT, CAC
Isoleucine	I	Ile	ATT, ATC, ATA
Lycine	K	Lys	AAA, AAG
Leucine	L	Leu	CTT, CTC, CTA, CTG, TTA, TTG
Methionine	M	Met	ATG
Asparagine	N	Asn	AAT, AAC
Proline	P	Pro	CCT, CCC, CCA, CCG
Glutamine	Q	Gln	CAA, CAG
Arginine	R	Arg	CGT, CGC, CGA, CGG, AGA, AGG
Serine	S	Ser	TCT, TCC, TCA, TCG, AGT, AGC
Threonine	T	Thr	ACT, ACC, ACA, ACG
Selenocysteine	U	Sec	Specific in-frame TGA codons[a]
Valine	V	Val	GTT, GTC, GTA, GTG
Tryptophan	W	Trp	TGG
Tyrosine	Y	Tyr	TAT, TAC
Glutamic acid or glutamine[b]	Z	Glx	See specific entries
Unknown/Other	X	Xaa	–

[a] Although TGA also functions as a stop codon, it is now recognized that certain in-frame TGA nucleotide sequences conditionally direct the incorporation of selenium as selenocysteine, the twenty-first amino acid.
[b] Used when identity is ambiguous. Once identity is confirmed, the appropriate amino acid-specific symbol is used. For glutamic acid/glutamine (Z, Glx), the symbols also refer to any substance that yields glutamic acid on acid hydrolysis of the peptide.

have been identified in neural tissue, many in extrahypothalamic tissue, where they act locally as neurotransmitters, neuromodulators, or trophic factors. Possibly related to their evolutionary history, many peptides have a longer duration of action than their amine counterparts. Thus, some have been described as modulating the tone of the information circuits over long time intervals. In other instances, the neuromodulatory action of peptides is to tune the gain of circuits in response to fast-acting, classic synaptic transmitters such as glutamate or γ-aminobutyric acid. Nevertheless, some peptides that meet the conventional criteria for neurotransmitters also signal information more dynamically. Although there is no consensus regarding a classification system for peptides, one approach has been to divide them according to their major anatomical location and effector mechanism. According to this scheme, biologically active peptides can be divided into the following classes: (1) gut peptides and central nervous system (CNS) neuropeptides, (2) hypothalamic neuropeptides that control the release of anterior pituitary hormones via a neurohormonal route, and (3) pituitary peptides that act primarily as hormones. Although this classification served a heuristic purpose, most peptides or homologous peptide gene products (e.g., gonadotropin-releasing hormone (GnRH) I and II and urocortins 1, 2, and 3) and their putative cognate receptors were subsequently identified in tissues other than those in which they were originally identified, in many cases acting in autocrine or paracrine, rather than endocrine, fashion. These tissues include the muscle; lungs; heart; adipose tissue; skin; and glial, immune, and hematopoietic cells and indicate the functional diversity of peptide action.

Biosynthesis and Processing of Peptides: The Proopiomelanocortin Model

Unlike biogenic amines, which are produced nonribosomally by biosynthetic enzymes, neuropeptides and hormones are derived from larger, mostly inactive precursor molecules that are gene products. Bioactive peptides are enzymatically cleaved and stored from these multipotential precursors. The biosynthesis and processing of one such precursor – proopiomelanocortin (POMC), the prohormone for adrenocorticotropic hormone (ACTH), melanocyte-stimulating hormone (MSH), lipotrophic hormone (LPH), and several opioids – has been well characterized and serves as a useful model for understanding the production of peptides *in vivo*. The POMC gene was the first neuropeptide precursor gene to be cloned and, like most peptide genes, has been found to be organized similarly across mammalian species. Spanning 6 kb, the POMC gene comprises three exons and two introns (see Figure 1). Upon transport from the nucleus to the cytoplasm, functional, mature POMC mRNA transcripts span 1100–1400 nt due to tissue-specific differences in the degree of 3′ polyadenylation of the mRNA. Transcripts subsequently bind to free ribosomes, where the translation of the N-terminus of the preprohormone begins. Exon 2 of the POMC gene codes for a 26-amino-acid signal (pre)peptide characteristic of all known secretory proteins. The binding of this α-helical signal peptide to a signal-recognition ribonucleoprotein and the association of this complex to a docking protein on the rough endoplasmic reticulum (RER) is required for further translation of peptide precursors. Once in the RER, the signal peptide is removed and the translation of the prohormone is completed. As can be seen in Figure 2, the translated POMC prohormone contains a multipotential N-terminal domain that contains the sequence for γ-MSH, a joining peptide, and domains that contain the full sequences for ACTH and β-lipotropin. From this point on, the prohormone undergoes posttranslational processing in a biologically regulated and tissue-specific manner to yield tissue- and state-specific peptide cocktails ready to be released from secretory vesicles. As can be seen in Figure 2, a multitude of peptides could be cleaved from this hormone, depending on physiological demands. Adding yet further flexibility to the system, certain peptide sequences are contained in multiple prohormones (e.g., met-enkephalin is present in both POMC and preproenkephalin). Much of this processing involves proteolytic cleavage by convertases at sites distal to dibasic residues (Arg-Lys, Lys-Arg, Lys-Lys) that flank the desired peptide sequence. For some peptides, a single basic residue (Arg) serves as the cleavage signal. Subsequently, the flanking dibasic peptides are snipped from the target peptide by carboxypeptidases. En route from the RER to the smooth endoplasmic reticulum, the Golgi apparatus and, ultimately, the secretory granules, the prohormone, and its peptide derivatives are also subject

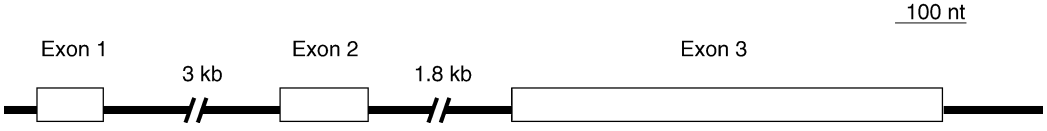


Figure 1 Representation of the structure of the rat proopiomelanocortin (POMC) gene.

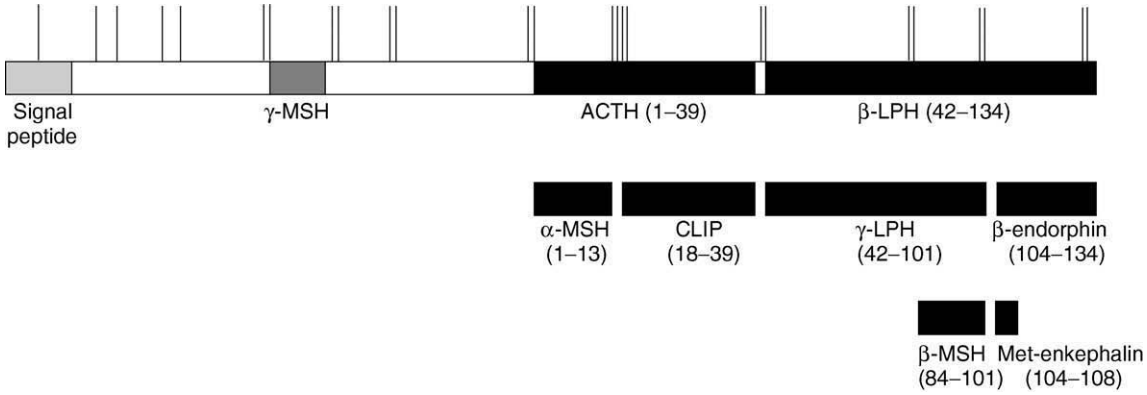


Figure 2 Representation of the structure of the bovine preproopiomelanocortin prohormone. The locations of known peptide components are shown in closed boxes with the amino acid numbers in parentheses. Paired basic amino acid residues (Arg-Lys, Lys-Arg, and Lys-Lys) and single cysteine residues, which represent possible cleavage sites, are marked with vertical bars. ACTH, adrenocorticotropic hormone; CLIP, peptide fragment; LPH, lipotrophic hormone; MSH, melanocyte-stimulating hormone.

to several enzymatic nonproteolytic modifications, including glycosylation, deglycosylation, acetylation, sulfation, phosphorylation, carboxymethylation, and C-terminal amidation. These modifications facilitate transport or storage, increase metabolic stability, and alter the product's bioactivity (e.g., the importance of C-terminal amidation for many peptides or of Ser3-acetylation for ghrelin). Thus, the process by which a particular peptide or peptide cocktail is localized in a secretory granule is extremely complex, with regulation (or dysregulation) possible at the level of prohormone synthesis, enzymatic processing, transport, and storage.

Peptide Release

Some peptides, notably growth factors, follow a constitutive route of release whereby they proceed directly from the Golgi network to the cell surface, bypassing storage in secretory granules. More typical of stress-related peptides found in endocrine cells and neurons, however, is a regulated route of release, as described in the POMC model, whereby the peptide is packaged in large, dense, core secretory vesicles until liberated by the appropriate stimulation of the cell. Neuropeptide release, like that of biogenic amine neurotransmitters, is induced by cell depolarization in a calcium-dependent manner. However, unlike bioamines, which are released from smaller clear synaptic vesicles during low-frequency activity, neuropeptides are generally released from large, dense core vesicles only on high-frequency burst firing. Because peptides lack reuptake sites to replenish secretory stores, *de novo* synthesis of new peptide is the major mechanism for replacement after release, such that peptide release can often be inferred indirectly from a compensatory increase in peptide mRNA expression.

Peptide Receptors

Once released, peptide ligands must interact with their receptors to transmit information. Neuropeptide receptors are generally expressed at low densities (1–10 pmol/mg protein at highest concentrations). Prior to the advent of molecular biological techniques, this made them difficult to isolate, clone, and characterize, although much recent progress has been made. They have been observed both pre- and postsynaptically on the soma, axons, and dendrites of neurons, indicating a complex mode of communication.

G-protein-coupled receptors

With few exceptions, peptide receptors are guanine nucleotide-binding (G)-protein-coupled receptors (GPCRs). The G-protein-coupled superfamily of receptors, which represents approximately 80% of all known receptors, has a highly conserved domain structure, with seven predicted transmembrane α-helices. The hydrophobic transmembrane core regions show the greatest homology across peptide receptors, whereas variation in the intra- and extracellular domains yields the diversity of receptor families and subtypes. A conserved, paired acid–base residue at the base of the third transmembrane appears to play a role in G-protein activation. The full range of nonphotoreceptor G-protein-mediated signal transduction mechanisms has been described for these receptors, including the positive and negative regulation of adenylyl cyclase, stimulation of phospholipase C, and positive and negative regulation of ion channels (Ca^{2+}, K^+, and nonselective Na^+/K^+). Of approximately 360 GPCRs coded in the human genome and identified in a database maintained by the International Union of Pharmacology Basic and Clinical Pharmacology Committee on Receptor Nomenclature and

Drug Classification, approximately one-fourth are known to have peptide ligands. Of peptide GPCRs, approximately 70 belong to the largest family, the class A or rhodopsinlike receptors; 15 belong to the class B (secretinlike) family of receptors for large polypeptidelike hormones; and none belong to the class C receptor family. A further 70 orphan GPCRs are estimated to have peptide ligands.

A few peptide receptors belong to the guanylate cyclase (rather than GPCR) family of membrane receptors. Ligands for these receptors, which are single-transmembrane proteins that regulate intracellular cyclic guanine monophosphate (cGMP) accumulation, include atrial natriuretic peptide, C-type natriuretic peptide, uroguanylin, and guanylin. Finally, a very limited number of peptides bind to other receptor types, including an ionotropic receptor for an invertebrate tetrapeptide FMFRamide and sortilin, a large, single-transmembrane receptor for neurotensin that has a long extracellular domain.

Theories of Peptide–Receptor Interaction

Peptides' activation of receptors cannot be explained by a simple lock-and-key model of amine–receptor interaction. Rather, two heuristic models, although theoretical, provide a framework for understanding peptide–receptor interactions: a zipper model and an address–message model. The former emphasizes that the association of a single amino acid on the peptide to a complementary receptor binding site initiates a series of interactions that results in a receptor-bound activating confirmation. The latter notes that certain portions of the peptide (the address) are more important for binding the ligand with the receptor, whereas other portions (the message) are more important for initiating the confirmational changes in the receptor that lead to signal transduction. Many, if not most, peptide receptors also exhibit a state of constitutive activation in the absence of ligand. Like their ligands, peptide receptors undergo posttranslational processing, including alternative splicing and differential N-glycosylation and phosphorylation. These modifications, which alter receptor binding affinities and signal-transduction efficiency, form the basis of several receptor subtypes.

Desensitization and Regulation of Peptide Receptors

Many peptide receptors, like their amine counterparts, undergo rapid-onset (<30 min) homologous receptor desensitization. The C-terminal tail of the receptor appears to participate in this process, perhaps by modifying intracellular protein interactions. Some receptors also exhibit heterologous augmentation of homologous desensitization; for example, arginine vasopressin synergistically enhances corticotropin-releasing factor (CRF)-induced desensitization of CRF receptors. However, the mechanism for this process is unclear. Finally, peptide receptor synthesis and expression also are subject to homologous and heterologous regulation.

Modulators of Peptide Action

Unlike biogenic amines, whose actions are terminated in part by cellular reuptake, peptide action is modulated by two specific largely extracellular mechanisms – peptidases and peptide-binding proteins.

Peptidases

Proteases are enzymes encoded by 2% of genes in all living organisms that hydrolyze peptide bonds, potentially with ensuing functional consequences. Proteases can be classified as the somewhat more common endopeptidases, which cleave at internal residues, or as exopeptidases, which break terminal peptide bonds. Exopeptidases are further subclassified into aminopeptidases and carboxypeptidases, according to their localized site of action. Proteases also can be described according to their chemical site of action (e.g., cysteine, serine, or threonine proteases). Peptidases exhibit varying degrees of substrate specificity, with few exhibiting high selectivity (e.g., the thyrotropin-releasing hormone-degrading ectoenzyme specifically inactivates thyrotropin-releasing hormone). Many peptidases, such as endopeptidase 24.11, angiotensin-converting enzyme, and pyroglutamate aminopeptidase II, are synaptosomally located, poised to limit neurotransmission by inactivating peptides released into the synaptic cleft. Indeed, peptidase release appears to be induced by neuronal depolarization. Other peptidases, however, such as aminopeptidases A and M, dipeptidyl aminopeptidase IV, and carboxypeptidase, appear to be located in the microvasculature and to be situated to limit volume transmission, or signaling occurring over a great distance, as well as to limit passage across the blood–brain barrier. A third class of membrane-associated cytoplasmic peptidases in the neural and glial tissue, for example, pyroglutamyl peptidase I, also may help to inactivate the released peptide. Finally, receptor-mediated endocytosis, with subsequent lysosomal degradation, is an additional route of limiting peptide activity. Each peptide is subject to inactivation by a variety of peptidases, with each reaction yielding different metabolic products. Because many of these metabolic products can act as agonists

or antagonists (in a manner sometimes functionally distinct from those of the parent peptide – for example, the conversion of peptide YY to peptide $YY_{3–36}$), the physiological control of bioavailable peptidases is an additional important determinant of peptide action. Protease activity is regulated *in vivo* by endogenous inhibitors, zymogen activation, and the modulation of their synthesis or degradation. Because of peptidases, natural peptides generally have very short half-lives (on the order of minutes) *in vivo*. Peptides with deletions, unnatural amino acids, or substitutions that eliminate a natural cleavage site can result in compounds with greater metabolic stability. Depending on the specific intrinsic activity of the analog, it could then function as a long-acting antagonist (e.g., astressin-B compared to corticotropin-releasing factor), an agonist (exendin-4 compared to glucagon-like peptide-1) or an inhibitory superagonist ([D-Trp^6, Pro^9-NEt]-GnRH compared to GnRH). Similarly, inhibitors of peptidase action can augment or prolong peptide action. The MEROPS database maintains an integrated list of known peptidases and peptidase inhibitors classified according to their structural similarity or catalytic site of action.

Peptide-Binding Proteins

Peptide-binding proteins are soluble or membrane-bound proteins to which peptides bind in lieu of their cognate receptor. Often, they act to sequester peptide ligands, thereby limiting their bioavailability (e.g., CRF-binding protein). Some binding proteins (e.g., insulin-like growth factor-binding protein), however, potentiate peptide action by protecting peptides from extracellular peptidases and forming a bound complex that associates with the membrane, thereby facilitating peptide–receptor interaction. Binding proteins are physiologically regulated through enzymatic modifications of their affinity for the peptide ligand or through changes in synthesis and expression, with consequent effects on peptidergic activity.

Acknowledgments

Supported by NIH grants DK64871 and DK26741 from the National Institute of Diabetes and Digestive and Kidney Diseases. This is publication number 17220-NP from The Scripps Research Institute.

See Also the Following Articles

Beta-Endorphin; Corticotropin Releasing Factor (CRF); Endocrine Systems; Gastrointestinal Effects; Hypothalamic-Pituitary-Adrenal; Anatomy of the HPA Axis; Annexin; Opioids; Oxytocin; Pro-opiomelanocortin (POMC); Pituitary Regulation, Role of; Protein Synthesis; Secretagogue; Urocortins; Vasoactive Peptides; Vasopressin.

Further Reading

Crawley, J. N. and McLean, S. (eds.) (1996). Preface. In: *Special issue on neuropeptides: basic and clinical advances, Annals of the New York Academy of Sciences* **780**, ix–xi.

Guillemin, R. (2005). Hypothalamic hormones a.k.a. hypothalamic releasing factors. *Journal of Endocrinology* **184**, 11–28.

Harmar, A. J. (2004). Receptors for gut peptides. *Best Practice & Research: Clinical Endocrinology & Metabolism* **18**, 463–475.

Holly, J. (2004). Physiology of the IGF system. *Novartis Foundation Symposiums* **262**, 19–26.

Hökfelt, T., Broberger, C., Xu, Z-Q., et al. (2000). Neuropeptides – an overview. *Neuropharmacology* **39**, 1337–1356.

Rawlings, N. D., Tolle, D. P. and Barrett, A. J. (2004). MEROPS: the peptidase database. *Nucleic Acids Research* **32**, D160–D164.

Relevant Websites

International Union of Pharmacology Basic and Clinical Pharmacology Committee on Receptor Nomenclature and Drug Classification database. http://www.iuphar-db.org.

MEROPS database. http://merops.sanger.ac.uk.

Perinatal (Fetal) Stress/Barker's Hypothesis *See:* Fetal Stress.

Perinatal Dexamethasone

E Fuchs
German Primate Center and Medical School, University of Göttingen, Göttingen, Germany

© 2007 Elsevier Inc. All rights reserved.

Early Life Programming
Dexamethasone: Actions and Potencies
Placental Metabolism of Glucocorticoids
Dexamethasone and Fetal Development
Metabolic Programming
Effects on Brain, Endocrine, and Immune Functioning and Behavior
Intergenerational Effects of Dexamethasone
The Underlying Mechanism: A Permanent Alteration of Gene Expression?
Conclusions

Glossary

Dexamethasone A synthetic fluorinated glucocorticoid hormone structurally related to cortisol.

Early life programming Programming reflects the action of a factor during a sensitive developmental period, or time window, to affect the development and organization of specific tissues that are concurrently vulnerable, producing effects that persist throughout life. Different cells and tissues are sensitive at different times, so the effects of environmental challenges will have distinct effects, depending not only on the challenge involved but also on its timing.

Early Life Programming

Lifestyle and genetic predisposition are considered to be the main determinants of noncommunicable diseases in humans. However, increasing evidence from studies in several distinct human populations worldwide indicates that many disorders occurring in adult life date back to influences of the intrauterine environment during the prenatal period. Examples of illnesses for which there is evidence of fetal origin are hypertension, coronary heart disease, diabetes, and neurological disorders. Behavioral disorders during childhood and adolescence may also be programmed during the prenatal period. As it is generally accepted that a multitude of pre- and postnatal influences are involved in pathophysiological processes that predispose individuals to such diseases, it is plausible to propose that there are factors triggering fetal development, thus amplifying or minimizing certain long-term processes. One identified trigger is hormonal disturbance such as high glucocorticoid exposure during pregnancy. However, most of the studies in humans have been restricted to either epidemiological or some physiological measurements that in sum cannot explain the cause and basic mechanisms underlying the pathophysiological processes.

A high percentage of pregnant women with an increased risk of preterm delivery are treated with glucocorticoids for enhancement of fetal lung maturation before term. The underlying mechanism for why this treatment is used in fetal pulmonary development is the ability of glucocorticoids to turn off cell proliferation and initiate mature cell function. In addition to being used in cases of preterm labor, synthetic glucocorticoids are administered to pregnant women in other clinical situations, including congenital adrenal hyperplasia and a wide spectrum of autoimmune disorders such as multiple sclerosis, asthma, rheumatic diseases, or chronic inflammatory intestinal diseases (M. Crohn, colitis ulcerosa). Often these treatments occur in early pregnancy, when it has not yet been diagnosed, and in some cases throughout gestation. There is currently little information about the impact of such treatment regimes on organ development. Epidemiological data indicate that fetuses exposed to synthetic glucocorticoids such as dexamethasone display significant increases in emotionality, unsociability, and behavioral problems in later life as children. However, further resolution of the safety of these treatment regimes is urgently needed. As a result, two National Institutes of Health (NIH) consensus reports published in 1995 and 2001 suggested that repeated or prolonged treatment should be avoided during pregnancy and highlighted the requirement for further research on the effects of prenatal glucocorticoids on organ development and function after birth.

Dexamethasone: Actions and Potencies

Synthetic glucocorticoids were developed by modifying various sites on the cortisol molecule in order to, for example, enhance the anti-inflammatory and reduce the Na^+-retaining mineralocorticoid activity. Among synthetic agents, the fluorinated steroid dexamethasone (DEX) has the lowest mineralocorticoid effect; it is characterized by a very long half-life, with the consequence of a greater potential to cause glucocorticoid-related side effects. DEX can be administered orally and, as with other steroids, is rapidly absorbed from the jejunum with high bioavailability,

Table 1 Comparative potencies of cortisol and dexamethasone

	Equivalent dose (mg)	*Plasma half-life (min)*	*Duration of action (h)*	*Mineralocorticoid potency*	*Glucocorticoid potency*	*Relative placental inactivation (% of drug inactivated)*
Cortisol	20	90	8–12	1	1	68
Dexamethasone	0.5	240	36–72	0	30	2

From Lin and Paget (2002).

meaning that up to 90% of an orally administered dose appears in the circulation, although bioavailability after oral administration is less than after parenteral administration because of substantial first-pass liver metabolism (Table 1).

DEX is used in the differential diagnosis of hypercortisolemia as it occurs, for example, in Cushing's disease, atypical pituitary adenomas, or depression (see **Hypocortisolism and Stress**). In this context, it is interesting to note that at least the adult rodent brain appears to be protected against high amounts of DEX by a multidrug-resistant P-glycoprotein located in the blood–brain barrier.

Placental Metabolism of Glucocorticoids

Transplacental passage of the different corticosteroid molecules varies, so different substances have to be used for maternal than for fetal targets. Dexamethasone and betamethasone, a second member of the clinically used fluorinated glucocorticoids, cross the placenta more readily than the natural glucocorticoid cortisol does because they are not bound to transporter proteins in the maternal blood, nor are they as easily converted to cortisone by placental enzymes such as 11β-hydroxysteroid dehydrogenase (11β-HSD). The relative inactivation of various corticosteroids to their 11-ketometabolites, through action of placental 11β-HSD, ranges from 2% (DEX) to 68% (cortisol; see Table 1). Within the placenta, two isoforms of 11β-HSD exist: a low-affinity NADP-dependent dehydrogenase/oxoreductase and a high affinity NAD-dependent dehydrogenase. Expression and distribution of these isoforms within the placenta most likely have an impact on levels of active glucocorticoids in fetal blood so that a complex interplay of factors determines the degree of fetal exposition.

Dexamethasone and Fetal Development

The assumption that excessive fetal exposure to glucocorticoid hormones constitutes an origin of adult diseases is based on studies in experimental animals. Findings in laboratory rodents (mice, rats, guinea pigs) concur with observations in ruminants and monkeys, mainly rhesus monkeys, documenting many lasting endocrine, neural, and behavioral effects when immature animals are exposed to steroidal hormones early in life. To model perinatal hyperexposure to glucocorticoids under controlled conditions, many research groups have studied the *in utero* effects of DEX, whose cellular effects are mediated via the glucocorticoid receptors (GRs) that are present in nearly all cells of the body (see **Glucocorticoids, Effects of Stress on**). It is known that synergistic and/or antagonistic interactions between activated GRs and several transcription factors regulate gene expression. Many organs, such as the brain, and the cardiovascular, skeletal, immune, reproductive, and urogenital systems contain cells that are vulnerable to changes in glucocorticoid homeostasis, and it is very likely that excess antenatal glucocorticoids such as DEX have multiple pathophysiological effects. The developmental consequences of prenatal DEX treatment are summarized in Figure 1.

Animal studies have revealed that, via the effects on gene transcription, glucocorticoids exert well-documented prenatal programming actions in a variety of target tissues to induce effects that may persist throughout the life span. For example, the adult rat offspring of pregnancies treated with DEX exhibited lifelong elevated blood pressure and high levels of plasma glucose, insulin, and corticosterone, as well as altered behavior. Variations in the effectiveness of the fetoplacental enzymatic barrier to maternal glucocorticoids because of changes in 11β-hydroxysteroid dehydrogenase type 2 (11β-HSD2), which catalyzes the rapid inactivation of cortisol (or corticosterone in rats) to inert cortisone (11-dehydrocorticosterone), correlate directly with birth weight in rats and humans. Thus, a deficiency of the placental barrier with respect to inactivation of maternal glucocorticoids, or overexposure to the hormones, may lead to increased effects of the hormones on gene expression in the fetal cells, and thus to prenatal programming. This idea is supported by the finding that, comparable to the effects of DEX, inhibition of 11β-HSD in pregnant rats by treating them with carbenoxolone causes reduced birth weight and hypertension later in life. Interestingly, the effects of carbenoxolone require the presence of endogenous maternal glucocorticoids: the offspring of adrenalectomized pregnant rats were found to be protected from the effects of carbenoxolone with respect to birth weight and adult physiology. However, it has

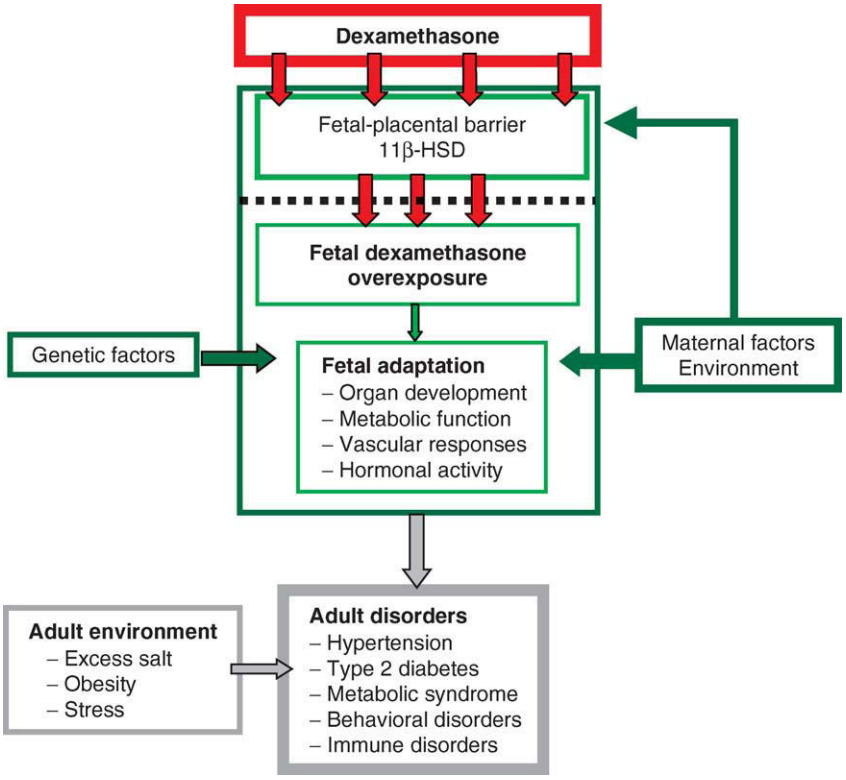


Figure 1 Developmental consequences of prenatal dexamethasone treatment. Fetal dexamethasone overexposure in combination with genetic and environmental factors will influence the development of many organ systems and programs tissue responses in the developing offspring, resulting in an increased risk for pathophysiological conditions later in life.

to be considered that carbenoxolone is a nonselective compound that inhibits both 11β-HSD isoforms as well as related dehydrogenases.

The timing of antenatal glucocorticoid exposure is important, in part because of the complex species-specific patterns of expression of the 11β-HSD isoenzymes that regulate maternal glucocorticoid transfer to the fetus. In rats, exposure to glucocorticoids during the last week of pregnancy is sufficient to produce permanent adult hypertension, whereas in sheep, the time window of sensitivity for such effects is earlier in gestation.

Mice and rats have been regarded as corticoid-sensitive species because their newborn pups are extremely immature at birth. Therefore, the relevance of rodent models of prenatal glucocorticoid overexposure has sometimes been questioned. However, similar types of adverse effects have also been observed in sheep, including an influence of physical growth and suppression of maternal and fetal adrenal activity. These glucocorticoid effects in sheep are surprising because high maternal cortisol levels play an important role in the normal physiology of their pregnancy and parturition. It is possible that in certain species the fetus is typically protected from excessive transplacental diffusion of cortisol/corticosterone through the enzyme 11β-HSD.

In monkeys, the natural ligand cortisol is metabolized even more rapidly in the fetus than in the full-term neonate. Conversely, DEX takes considerably longer to clear from the circulation. This extended half-life may be important when DEX is administered to the mother while the baby is still *in utero* and thus depends on the transfer of steroid hormones back across the placenta, rather than on its own renal excretion after birth.

Unfortunately, not enough animal studies have been conducted with betamethasone to permit a systematic comparison with DEX. In general, the research in humans tends to find more side effects with DEX than with betamethasone.

Metabolic Programming

The preceding findings imply that glucocorticoids exert organizational effects on the immature system during defined developmental stages, leading to changes in physiology that have persistent consequences throughout the lifetime. Because reduced prenatal growth and low birth weight are determined largely by nongenetic factors, these findings have led to the fetal origin hypothesis. It proposes that the physiology of the fetus adapts to the adverse intrauterine environment, which in turn reduces fetal growth

and programs the lifelong changes. This programming predisposes an individual to the metabolic syndrome that manifests itself only in adulthood. It emanates from a general disturbance of the body's metabolic processes and leads to heart disease, diabetes, and kidney failure. These illnesses have long been observed as being somehow linked to modern civilization, but the fact that they constitute an entire syndrome has been recognized only recently.

Blood Pressure

The impact of prenatal glucocorticoid exposure during development on arterial blood pressure has been investigated in the offspring of several species. Treatment of pregnant rats with DEX results in persistent elevations of arterial blood pressure in the adult offspring. Maternal administration of the 11β-HSD inhibitor carbenoxolone leads to elevated blood pressure in the adult offspring. As previously mentioned, this effect requires the presence of maternal glucocorticoids, as the offspring of adrenalectomized pregnant rats are protected from the adverse actions of carbenoxolone.

The mechanisms of glucocorticoid-programmed adult hypertension are not clearly understood but probably involve a variety of processes. Potential factors may involve primary effects on the heart by interfering with the development of cardiac noradrenergic innervation, modifications in vascular responsivity, alterations in renal structure (irreversible reductions in nephron number), and enhanced activity of the renin-angiotensin system.

Glucose Homeostasis

Prenatal overexposure to glucocorticoids also causes permanent hyperglycemia and hyperinsulinemia in the adult offspring of rats and sheep. As with blood pressure, the window of sensitivity for programming hyperglycemia in the rat occurs in the last third of gestation. Earlier DEX exposure or *post partum* treatments do not program hyperglycemia/hyperinsulinemia, suggesting that there is a narrow window for this effect. The mechanisms through which prenatal superphysiological levels of glucocorticoids program hyperglycemia have not been fully determined but may involve derangements in several target organs such as the liver, the pancreas, fat, and skeletal muscle.

Glucocorticoids regulate the expression of several hepatic metabolic enzymes, including phosphoenolpyruvate carboxykinase (PEPCK). This enzyme develops in late gestation and catalyzes a rate-limiting step in gluconeogenesis. In rats, exposure to excess DEX *in utero* leads to offspring with permanent elevations in PEPCK mRNA and enzyme activity. Interestingly the PEPCK activity is increased in patients with type 2 diabetes.

Relatively little is known about the effects of prenatal exposure to glucocorticoids on the development of the endocrine functions of the pancreas. However, indirect evidence suggests that excessive prenatal glucocorticoid levels may cause long-term beta cell dysfunction.

Prenatal DEX exposure in rats programs fat metabolism. Adult offspring of dams exposed to DEX in pregnancy have increased intra-abdominal fat depots, with a parallel increase in leptin levels. The same antenatal treatment may also affect skeletal muscle glucose metabolism.

Effects on Brain, Endocrine, and Immune Functioning and Behavior

In addition to their metabolic role, glucocorticoids are essential for normal brain development. Glucocorticoids are important for normal maturation in most regions of the developing central nervous system, initiating terminal maturation and remodeling axons and dendrites, and for cell survival.

Exposure to glucocorticoids *in utero* has widespread acute effects on neuronal structure and synapse formation and may permanently alter brain structure. Prenatal glucocorticoid administration retards brain weight at birth in sheep, delaying maturation of neurons, myelination, glia, and vasculature. Rhesus monkeys that were prenatally exposed to DEX exhibited marked atrophy of hippocampal cells at birth, and their hippocampi failed to recover to normal size, even at 2 years *post partum*. Human and animal studies have demonstrated that altered hippocampal structure may be associated with memory impairment and behavioral alterations (see **Hippocampus, Corticosteroid Effects on**).

Overall, the development of the hippocampus in rhesus monkeys is comparable to that in humans, although it is difficult to specifically compare the rate of fetal growth and stage of maturity at birth. However, in humans, important phases in maturation of the brain, including brain growth, take place during the postpartum period. At birth, the brain of a rhesus monkey is 60% of its adult size, whereas the human brain is only 24% of its adult size. Thus, the impact of prenatal DEX overexposure on the developing nervous system may be of a greater magnitude in the monkey.

The hypothalamic-pituitary-adrenocortical (HPA) circuit, which is central for the integration of an individual's endocrine and behavioral responses and for the whole body's homeostasis, appears highly sensitive to excess glucocorticoids during development. A number of animal studies have shown that exposure to synthetic glucocorticoids *in utero* results in disturbed hypothalamic-pituitary circuit functioning.

Studies in rats provided evidence that perinatal DEX treatment enhances susceptibility to experimental autoimmune encephalitis (EAE) in adult life. EAE is an experimental autoimmune inflammatory disease that serves as an animal model for multiple sclerosis. Low reactivity of the HPA axis and increased Th1-type cytokine production are discussed as potential mechanisms for this increased susceptibility in adult rats after DEX treatment. Furthermore, perinatally DEX-treated rats have a long-term decreased production of cytokines by macrophages. This may lead to increased susceptibility to bacterial infections in later life. Apparently, perinatal DEX treatment has permanent programming effects on endocrine as well as on immune functioning in adult life.

As revealed mainly by studies in rats, overexposure to glucocorticoids *in utero* leads to alterations in adult behavior. Late gestational DEX in rats apparently impairs coping in aversive situations later in life. Behavioral changes in adult rats exposed prenatally to glucocorticoids may be associated with altered functioning of the amygdala, a key brain structure in the regulation of fear and anxiety. Interestingly, in rats, adult offspring of DEX-treated dams show increased corticotropin-releasing hormone levels, specifically in the central nucleus of the amygdala.

Intergenerational Effects of Dexamethasone

The effects of fetal programming may not be limited to the first-generation offspring. A recent study investigated epigenetic effects in the DEX-programmed rat, a model in which fetal exposure to excess corticosteroids results in low birth weight, adult glucose intolerance, and increased activity of the key hepatic gluconeogenic enzyme PEPCK. Male offspring of female rats that had been exposed prenatally to DEX, but that were not treated during their pregnancy, also had reduced birth weight, glucose intolerance, and elevated hepatic PEPCK activity. These effects resolved in a third generation. Similar lifelong programming was observed in the offspring of male rats exposed prenatally to DEX mated with control females. The persistence of such programming effects across several generations, transmitted by either maternal or paternal lines, indicates – at least in rodents – the potential importance of non-genomic factors in the intergenerational inheritance of the programming phenotype.

The Underlying Mechanism: A Permanent Alteration of Gene Expression?

The GR is critical for cell function, and GR gene expression shows marked tissue-specific regulation. In the peripheral tissues, such as the liver and visceral adipose tissue, prenatal glucocorticoid exposure causes a permanent increase in GR expression; this upregulation is paralleled by increased glucocorticoid sensitivity in these tissues.

Until recently it was unclear how perinatal glucocorticoid overexposure could permanently alter gene expression. Recent evidence suggests that this priming may involve selective methylation/demethylation processes of specific promoters of the GR gene. Changes in the GR promoter DNA methylation pattern were associated with altered histone acetylation and transcription factor (NGFI-A) binding to the GR promoter. Treatment of the adult rat offspring with a histone deacetylase inhibitor removed the epigenetic differences in histone acetylation and DNA methylation and changes in NGFI-A binding. During development, such target promoter demethylation occurs before birth and may tune the promoter to remember regulatory events occurring during development. This novel mechanism of gene control by environmental events early in life, which then persists throughout the life span, needs to be confirmed in species other than rodents.

Conclusions

Balanced levels of glucocorticoids are of central importance for normal development, and it can be assumed that glucocorticoid overexposure during critical periods of fetal development may have long-term effects that could be deleterious later in life. To date, most of the studies in humans have been restricted to either epidemiological or some physiological measurements that in sum cannot explain the cause and basic mechanisms underlying the pathophysiological processes. In contrast, animal studies have indicated a link between glucocorticoid overexposure and an increased risk for pathophysiological conditions later in life. Until now, experimental studies on the mechanisms induced by prenatal hyperexposure to glucocorticoids such as DEX have been conducted predominantly in rodents. However, rodents are far from ideal models of primate pregnancy due to their multiple offspring and short gestation lengths and life span. Clearly, mechanistic studies need to be extended to species more closely related to humans.

See Also the Following Articles

Hypocortisolism and Stress; Hippocampus, Corticosteroid Effects on; Glucocorticoids, Effects of Stress on.

Further Reading

Aghajafari, F., Murphy, K., Matthews, S., et al. (2002). Repeated doses of antenatal corticosteroids in animals:

a systematic review. *American Journal of Obstetrics and Gynecology* **186**, 843–849.
Bakker, J. M., Kavelaars, A., Kamphuis, P. J., et al. (2000). Neonatal dexamethasone treatment increases susceptibility to experimental autoimmune disease in adult rats. *Journal of Immunology* **165**, 5932–5937.
Bakker, J. M., van Bel, F. and Heijnen, C. J. (2001). Neonatal glucocorticoids and the developing brain: short-term treatment with life-long consequences? *Trends in Neuroscience* **24**, 649–653.
Berger, B., Alvarez, C. and Goldman-Rakic, P. S. (1993). Neurochemical development of the hippocampal region in the fetal rhesus monkey. I. Early appearance of peptides, calcium-binding proteins, DARPP-32, and monoamine innervation in the entorhinal cortex during the first half of gestation (E47–90). *Hippocampus* **3**, 279–305.
Coe, C. L. and Lubach, G. R. (2005). Developmental consequences of antenatal dexamethasone treatment in nonhuman primates. *Neuroscience and Biobehavioral Reviews* **29**, 227–235.
Drake, A. J., Walker, B. R. and Seckl, J. R. (2005). Intergenerational consequences of fetal programming by *in utero* exposure to glucocorticoids in rats. *American Journal of Physiology: Regulatory, Integrative, and Comparative Physiology* **288**, R34–R38.
Kay, H. H., Bird, I. M., Coe, C. L. and Dudley, D. J. (2000). Antenatal steroid treatment and adverse fetal effects. What is the evidence? *Journal of Society for Gynecologic Investigation* **7**, 269–278.
Krozowski, Z., Li, K. X., Koyama, K., et al. (1999). The type I and type II 11beta-hydroxysteroid dehydrogenase enzymes. *Journal of Steroid Biochemistry and Molecular Biology* **69**, 391–401.
Lin, A. N. and Paget, S. A. (eds.) (2002). *Principles of corticosteroid therapy*. London: Arnold.
Matthews, S. G. (2000). Antenatal glucocorticoids and programming of the developing CNS. *Pediatric Research* **47**, 291–300.
Meijer, O. C., de Lange, E. C., Breimer, D. D., et al. (1998). Penetration of dexamethasone into brain glucocorticoid targets is enhanced in mdr1A P-glycoprotein knockout mice. *Endocrinology* **139**, 1789–1793.
National Institutes of Health Consensus Development Conference Statement. August 17–18, 2000 (2001). Antenatal corticosteroids revisited: repeat courses. *Obstetrics and Gynecology* **98**, 144–150.
NIH Consensus Developmental Panel on the Effect of Corticosteroids for Fetal Maturation on Perinatal Outcomes. (1995). Effect of corticosteroids for fetal maturation on perinatal outcomes. *Journal of the American Medical Association* **273**, 413–418.
Seckl, J. R. (2004). Prenatal glucocorticoids and long-term programming. *European Journal of Endocrinology* **151**, 1–16.
Uno, H., Eisele, S., Sakai, A., et al. (1994). Neurotoxicity of glucocorticoids in the primate brain. *Hormones and Behavior* **28**, 336–348.

Perinatal Stress Animal Studies *See:* Fetal Stress.

Persian Gulf War, Stress Effects of

S M Southwick, D Vojvoda, C A Morgan III and D Lipschitz
Yale University School of Medicine, New Haven, CT, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 142–148, © 2000 Elsevier Inc.

Health Risks of the Gulf War
Mental Health
Predictors of Posttraumatic Stress Disorder
Physical Symptoms in Gulf War Veterans
Relationship between Physical and Psychological Complaints
Summary

Glossary

Posttraumatic stress disorder (PTSD)	This disorder results from situations in which an individual believes that he or she or someone else is in danger of being seriously injured or killed. Symptoms of PTSD include repetitive re-experiencing of the trauma in the form of intrusive memories, nightmares, and flashbacks; persistent avoidance of stimuli reminiscent of the trauma; a general numbing of emotions; and increased arousal as expressed by hypervigilance, exaggerated startle, irritability, and insomnia.

Health Risks of the Gulf War

In 1990 the United States and allied countries responded to the Iraqi invasion of Kuwait by deploying

a large number of soldiers to the Persian Gulf. Operation Desert Storm began in 1991 when Allied forces launched a 5-week intensive air attack on Iraq. This was followed by a 4-day ground war. The number of deaths to American and Allied forces during the Gulf war was limited compared to World War I and II, the Korean war, and the Vietnam war. Nevertheless, since the end of the Gulf war, a substantial number of Desert Storm veterans have experienced a host of physical and psychological symptoms that, in some cases, have had a pronounced impact on psychosocial functioning.

Military personnel stationed in the Gulf were exposed to a wide variety of potential health risks. These included a hostile environment (e.g., extreme temperatures, humidity and rainfall, blowing sand, crowded living conditions), infectious agents (e.g., leichmaniasis, sandfly fever), medical prophylaxis (e.g., pyridostigmine bromide, anthrax, and botulinum toxin vaccines), occupational exposures (e.g., chemical exposures associated with routine maintenance and repair of equipment, fuels), environmental hazards (e.g., pesticides, smoke from oil well fires), depleted uranium (used in some artillery shells), the possibility of chemical and biological warfare agents (e.g., organophosphate nerve agents, sulfur mustard, anthrax, botulinus toxin), and repetitive physiological and psychological stress.

Soldiers serving in the Persian Gulf were exposed to numerous potential psychosocial stressors. Stressors associated with deployment included uncertainty about parameters of deployment (e.g., the period of waiting for deployment, unknown length of tour and return date), overestimation of Iraqi military forces, media anticipation of high casualty rates, and separation from home and family. As in all wars, soldiers described the anticipation of combat as being highly stressful. In a study of 1524 deployed military personnel, 80% of subjects reported feeling stressed by the persistent threat of SCUD missile attacks, terrorist attacks, and biological and chemical warfare. Many experienced strong emotional responses to repeated missile warnings and the rapid donning of protective gear that they did not fully trust would be effective against all possible toxic agents.

Reports of directly experienced traumatic events vary depending on the units studied. A series of scientific investigations commissioned by the U.S. Department of Defense found that 40% of soldiers in the units studied experienced incoming indirect fire, between 40 and 70% reported exposure to mines and booby traps, 40% in two units recalled firing rounds at the enemy, and 80% in one unit and 60% in another claimed to see enemy killed or wounded in action. In two units more than 20% of soldiers believed they were in imminent danger of being killed at some point during their tour in the Persian Gulf. Reports of seeing an American wounded by friendly fire had great emotional impact, as did work in grave registration units. Sadly, some traumas experienced in the Gulf, such as sexual harassment and assault, were inflicted by Americans on their own fellow soldiers.

Other frequently cited Gulf war stressors include austere physical environment and severe climatic conditions, heavy workload, sleep deprivation, inadequate supplies and equipment, lack of information, poor leadership and unit morale, and unsatisfactory living conditions.

Mental Health

The psychological consequences of war have been described for centuries and have been referred to by a variety of different terms, including shell shock, irritable heart of soldiers, physioneurosis, and, most recently, posttraumatic stress disorder (PTSD). PTSD results from situations in which the individual believes that he or she or someone else is in danger of being seriously injured or killed. Symptoms of PTSD include repetitive re-experiencing of the trauma in the form of intrusive memories, nightmares, and flashbacks; persistent avoidance of stimuli reminiscent of the trauma; a general numbing of emotions; and increased arousal as expressed by hypervigilance, exaggerated startle, irritability, and insomnia. Depression, anxiety, substance use, and changes in personality traits are other frequently cited consequences of trauma.

Psychological Symptoms in Persian Gulf Veterans

Prior to the ground war it was estimated that acute psychiatric casualties might reach as high as 50 000. This estimate was based on experience from previous wars. However, during the Gulf war there appeared to be very few psychiatric casualties. Between August 1990 and August 1991, 476 soldiers were evacuated from southwest Asia for psychiatric reasons. This number constituted only 6.5% of all medical evacuations from the Persian Gulf. Similarly, only 158 soldiers from the U.S. Army 7th Corps received mental health treatment while stationed in the Gulf, and 99% were returned to duty after brief treatment. These low rates led to the belief that psychiatric problems in Gulf veterans would be lower than expected. Explanations for low rates included the brief nature and success of the war, limited exposure to traditional combat stressors, and high levels of perceived social support.

Since the end of the Gulf war, however, a number of different research groups have reported rates of psychiatric symptoms and disorders in Persian Gulf veterans that are high in light of the earlier evacuation

figures. The Fort Devons Operation Desert Storm Reunion Survey was conducted within 5 days of return to the United States before soldiers rejoined their families. The most commonly endorsed war zone stressors were formal alert for chemical or biological attack (approximately 75%), receiving incoming fire from large arms (approximately 73%), and witnessing death or disfigurement of enemy troops (approximately 48%). Four percent of 2136 men and 9% of 208 females met a presumptive diagnosis of PTSD. Symptoms of hyperarousal and exaggerated startle were the most commonly endorsed PTSD symptoms. Twenty-eight to 31% had general severity index (GSI) scores on the brief symptom inventory (BSI), which reflects general psychological stress, in the clinically significant range.

Approximately 6 months after Operation Desert Storm, Labatte and Snow administered a self-report questionnaire to 57 soldiers stationed in Germany who were members of a single mechanized infantry unit that suffered four casualties during the brief ground war. Fourteen soldiers (25%) reported traumatic nightmares within 2 months of the war. Nightmares began 2 to 6 months after the war for an additional nine soldiers (16%).

Two other studies assessed PTSD symptoms and general psychopathology in a large number of reservists within the first year of their return from the Persian Gulf. In a study of 591 Army, Navy, and Marine reservists, Perconte and co-workers found that those who had been deployed to the Gulf war combat zone had significantly higher Mississippi PTSD scores, higher Beck depression scores, and higher levels of global distress as measured by the symptom check list-90 (SCL-90) than reservists who had been activated but not deployed to combat zones. Over 15% of deployed reservists met Mississippi cutoff scores for PTSD compared to less than 4% of nondeployed reservists. Sutker and colleagues reported low levels of psychological distress in their overall group of 775 combat-deployed Reserve and National Guard troops. However, 12.5% of the sample were diagnosed with PTSD.

Approximately 1 year after the war, Holmes and colleagues evaluated 179 members of an Air National Guard unit who were never deployed to the Persian Gulf and compared them to 296 members of the same unit who were deployed. Most members of the deployed group did not experience any traditional combat stressors, although they were exposed to nontraditional combat stressors such as anticipation of missile attacks and biological warfare. Compared to the nondeployed, the deployed group had statistically higher mean scores on the Mississippi PTSD scale and on the GSI as well as on the obsessive-compulsive, depression, hostility, phobic anxiety, and paranoia subscales of the BSI; 6.8% of deployed and 1.7% of nondeployed were identified as having increased risk for PTSD. The findings of this study suggest that multiple noncombat stressors were distressing enough to cause both PTSD and non-PTSD psychological symptoms in some deployed veterans.

In a follow-up study of the Fort Devons sample 18–24 months after the war, nearly half of the sample reported no physical or psychological symptoms. Twenty-seven percent believed that their psychological health had deteriorated since the war. PTSD symptomatology was significantly associated with high health symptom endorsement. Twenty-two percent of soldiers with high combat exposure met a presumptive diagnosis of PTSD compared to 8.1% with low combat exposure. Data were consistent with a number of recent studies suggesting that an apparent association between combat exposure and health is significantly mediated by a diagnosis of PTSD.

In a survey administered approximately 2 years after the war to 1524 deployed and 2727 nondeployed military personnel from Hawaii and Pennsylvania, Stretch and co-workers reported possible PTSD in 8.0% of active duty veterans and in 9.35% of reserve veterans who had been deployed to the Persian Gulf. By comparison, 1.3% of active duty nondeployed and 2.1% of reservist nondeployed met possible criteria for PTSD. The sample was also compared to a large group of deployed active duty Army veterans from the XVIIIth and VIIth Corps. The possible risk for PTSD in these groups was 15.2 and 12.9%, respectively. The authors speculate that these units had higher rates of possible PTSD because their duty was more dangerous and because they were evaluated 1 year earlier. Deployed personnel also scored significantly higher on the somatization, obsessive-compulsive, and hostility subscales of the BSI and the GSI.

As part of a follow along study of Gulf veterans from two National Guard units, Southwick and colleagues reported a rate of presumptive PTSD in 8.3% of 84 subjects 6 months after the war and in 10% of 64 subjects 2 years after the war. The greatest increase in overall PTSD symptoms occurred between 1 month and 6 months after the war, and symptoms of hyperarousal were elevated above re-experiencing and avoidance symptoms at 1-month, 6-month, and 2-year evaluations. Morgan and colleagues found evidence for anniversary reactions in this same group of veterans.

Approximately 5 years after the war, the Iowa Persian Gulf study group conducted telephone interviews with 1896 Gulf war veterans and 1799 personnel who served in the military during the war but who were not stationed in the Persian Gulf. Persian Gulf war personnel reported higher rates of PTSD (1.9%

vs. 0.8%), depression (17.0% vs. 10.9%), cognitive dysfunction (18.7% vs. 7.6%), asthma (7.2% vs. 4.1%), fibromyalgia (19.2% vs. 9.6%), alcohol abuse (17.4% vs. 12.6%), chronic fatigue (1.3% vs. 0.3%), and sexual dysfunction compared to non-Persian Gulf war military personnel. Symptom levels were generally higher in Persian Gulf war National Guard reserve personnel than in Persian Gulf war regular military personnel.

Finally, in a report of 18 495 Gulf war veterans who presented for evaluation to the Department of Defense Comprehensive Clinical Evaluation Program, 36% met ICD-9-CM diagnoses for psychological conditions. In most cases, symptoms did not develop until after the war. There was no apparent association between individual physical or psychological symptoms and specific exposures. Compared with a referent group of patients having no symptoms, an excess prevalence of disorders was highest for psychological conditions compared to a variety of physical conditions such as digestive and skin conditions.

Psychological Symptoms in Civilians Affected by the Persian Gulf War

Clearly the Gulf war had profound stress-related effects on Iraqi military personnel and civilians living in Iraq and surrounding Middle Eastern countries. In general, the rates of death, physical injury, and psychological injury have not been reported accurately in the scientific literature. However, there have been a number of publications related to the effects of the Gulf war on American and Israeli civilians as well as Vietnam veterans and Holocaust survivors.

In American civilians, it has been reported that the Gulf war was associated with mild to moderate symptoms of anxiety and depression, particularly in those who had a relationship with a deployed soldier. In children of deployed American soldiers, a major predictor of psychological symptoms was level of symptoms in other household members. Although deployment rarely provoked pathological levels of symptoms in healthy children, it was associated with increases in symptoms of depression, especially among younger children. In a study of Israeli civilians that was conducted during the war, Solomon and co-workers found that 80% of individuals whose homes had been destroyed by Iraqi missiles described symptoms consistent with DSM-III-R criteria for PTSD. Similarly, Laor and colleagues reported elevated levels of Gulf war stress-related symptoms in children and mothers who were displaced from their homes due to damage from SCUD missiles and in young children living in families characterized by inadequate cohesion.

In Vietnam veterans, no consistent response to the Gulf war has been described. Reported responses have included support for the war, anger at the U.S. government for starting another war, irritability, intrusive thoughts, and increased depression and suicidality. In a study of 76 female Vietnam veterans with PTSD, Wolfe and colleagues reported some exacerbation of symptoms in most subjects, with the greatest increases in those who had high levels of preexisting PTSD symptoms. In a study of Holocaust survivors living in Israel during the Gulf war, Robinson and colleagues reported that many were still vulnerable 50 years after WWII and reported a revival of feelings and memories associated with the Holocaust.

Predictors of Posttraumatic Stress Disorder

Numerous investigations of WWII, Korean, and Vietnam veterans have shown that not all individuals develop PTSD and/or other psychological disorders after life-threatening traumas. Given an equal degree of traumatic exposure, some individuals are more likely than others to develop subsequent trauma-related psychopathology. In order to explain such individual differences in veterans of the Gulf war, researchers have studied a number of potential risk factors or predictors for the development of PTSD. These have included characteristics of the traumatic event, as well as available personal and environmental resources.

The most commonly cited risk factor for the development of PTSD is the nature and severity of the trauma itself. Many war-related investigations, including a number of Desert Storm studies, have demonstrated a clear dose–response relationship between the severity of psychological trauma and subsequent psychopathology (i.e., the more severe the trauma, the greater the likelihood of developing PTSD and/or other mental disorders). When traumatic exposure involves actual physical injury, the risk for subsequent psychopathology is even greater. In a hospital chart review study of medical and surgical patients who were evacuated from the Gulf to the Walter Reed Hospital, Brandt and co-workers reported a significantly higher rate of axis I disorders (and psychiatric symptoms of concern) in the hospital records of soldiers with traumatic injuries (e.g., shrapnel, gunshot wounds, motor vehicle accidents, SCUD missile attack) compared to soldiers without traumatic injuries (e.g., gastrointestinal bleeding, chest pain, malignancies). Other trauma-related risk factors identified in Gulf war veterans include exposure to dead or dying bodies, having a buddy wounded or killed in action, exposure to American soldiers killed or wounded by friendly fire, being in the reserve (e.g., lack of psychological preparation for deployment), and handling of human remains.

Personal variables that have been studied in Gulf war veterans include gender, race, intelligence, personality, hardiness, coping style, and perceived social

and family support. In a study of 653 Persian Gulf war zone-exposed veterans, Sutker and co-workers reported significantly greater psychological distress and PTSD symptoms in minority men compared to white men but no differences between war zone males and females. In a separate report, Sutker and colleagues found more avoidance, wishful thinking, and self-blame coping strategies; fewer problem-oriented coping strategies; lower scores on hardiness dimensions of commitment, control, and challenge; and less perceived social and family support among Gulf war veterans with PTSD ($n = 97$) compared to those without psychopathology ($n = 484$). No differences in intellectual sophistication as measured by the Shipley Institute of Living Scale were found between the two groups.

Physical Symptoms in Gulf War Veterans

Since the end of the Gulf war, a large number of veterans have reported an array of troubling physical symptoms, including fatigue, joint pain, gastrointestinal problems, skin rash, headache, memory loss, insomnia, and shortness of breath. In a relatively small percentage of veterans, specific etiologies for physical complaints have been linked directly to military service in the Gulf (e.g., leishmaniasis, malaria, and embedded fragments of uranium). However, for the majority of veterans with physical complaints, a specific relationship between physical symptoms and biological, chemical, or other environmental exposures has not been reported in the literature. Further, no specific medical diagnoses appear to explain most of the complaints. As a result, some have postulated a Gulf war syndrome to account for these unexplained physical symptoms.

While some studies clearly report a greater number of physical complaints among military personnel who served in the Gulf compared to those who did not serve in the Gulf, most have not found high rates of any one medical disorder or conclusive evidence for a new syndrome. To date, a number of expert panels (Defense Science Board Task Force, National Institutes of Health Technology, Institute of Medicine, Presidential Advisory Committee on Gulf War Veteran's Illnesses) have failed to identify a new Gulf war syndrome. For example, in 1996 the Institute of Medicine reported that "the committee has not identified scientific evidence to date demonstrating adverse health consequences linked with PGW service other than the documented incidents of leishmaniasis, combat-related or injury-related mortality or morbidity, and increased risk of psychiatric sequelae of deployment." Similarly, in 1996 Gray and associates reported no increased risk of prewar or postwar hospitalization for any cause among 547 076 Persian Gulf war veterans compared to 618 335 nondeployed. Persian Gulf war veterans did have higher rates of psychiatric hospitalization that appeared to be related to the use of drugs and alcohol. Two large-scale mortality studies have reported no increased mortality as a result of medical diseases in Persian Gulf war veterans. However, an excess number of deaths due to accidents and unintentional injury have been reported.

Relationship between Physical and Psychological Complaints

In a comprehensive literature review on the relationship among psychological trauma, posttraumatic stress disorder, and physical health in diverse trauma populations, Friedman and Schnurr reported that exposure to catastrophic events (independent of actual physical injury) was frequently associated with adverse health reports and increased medical utilization, morbidity, and mortality. Based on work by Wolfe and others, Friedman and Schnurr's review suggested that PTSD – the reaction to a trauma – may serve as an important mediator in the association between traumatic exposure and perceived physical health. Baker and co-workers found support for this hypothesis in a sample of 188 Desert Storm veterans where fatigue, nausea, muscle aches, dizziness, back pain, stomach ache, and numbness were much more likely to be reported by Persian Gulf war veterans with PTSD compared to those without PTSD. Similarly, Stretch and colleagues reported a probable link between Gulf war syndrome and PTSD, even though they found the relationship between stress and physical health to be relatively weak in a large group of Desert Storm veterans. Approximately 30% of veterans with Gulf war syndrome also met criteria for PTSD. In a study of female veterans conducted within 1 year of their return from the Gulf, Wolfe and colleagues found anxiety and PTSD, but not combat exposure, to be significant predictors of health symptoms.

Neurobiological, psychological, and behavioral hypotheses have been offered to explain the relationship between PTSD and physical health. As noted by Friedman and Schnurr, medical problems in trauma survivors may result from neurobiological alterations characteristic of PTSD, such as exaggerated cardiovascular reactivity, adrenergic dysregulation, disturbed sleep physiology, altered hypothalamic-pituitary-adrenal axis activity, and enhanced thyroid function. It is also possible that these neurobiological alterations increase susceptibility to immunologic disorders and infections. Psychological and behavioral factors that may contribute to health problems and that are commonly seen in survivors with PTSD include hostility,

depression, alcohol and drug use, smoking, poor social support, and avoidant and emotion-focused coping.

Summary

Soldiers serving in the Gulf war were exposed to a wide variety of potential health risks, including traumatic stress. Across multiple studies it appears that between 6 and 10% of Gulf war veterans developed PTSD as a result of their service in the Middle East. A host of other psychological and physical symptoms have been reported, and some researchers have proposed a Gulf war syndrome. In Gulf war veterans the precise relationship among stress, the neurobiological alterations that accompany stress, and toxic or infectious exposure currently is poorly understood and the subject of intense scientific investigation.

Acknowledgments

This work was partially supported by grant number DAMD17-96-1-6125, U.S. Army, and by the National Center for PTSD, West Haven, CT.

See Also the Following Articles

Combat Stress Reaction; Gulf War Syndrome, Psychological and Chemical Stressors; Korean Conflict, Stress Effects of; Posttraumatic Therapy; Vietnam Veterans, Postwar Experiences and Health Outcomes; War-Related Posttraumatic Stress Disorder, Treatment of.

Further Reading

Baker, D. G., Mendenhall, C. L., Simbartl, M. S., Magan, L. K. and Steinberg, J. L. (1997). Relationship between posttraumatic stress disorder and self-reported physical symptoms in Persian Gulf War veterans. *Archives of Internal Medicine* **157**, 2076–2078.

Brandt, G. T., Norwood, A. E., Ursano, R. J., et al. (1997). Psychiatric morbidity in medical and surgical patients evacuated from the Persian Gulf war. *Psychiatric Services* **48**, 102–104.

Friedman, M. J. and Schnurr, P. P. (1995). The relationship between trauma, posttraumatic stress disorder and physical health. In: Friedman, M. J., Charney, D. S. & Deutch, A. (eds.) *Neurobiological and clinical consequences of stress: from normal adaptation to PTSD*. Philadelphia, PA: Lippincott-Raven.

Gray, G. C., Coate, B. D., Anderson, C. M., et al. (1996). The postwar hospitalization experience of U.S. veterans of the Persian Gulf War. *New England Journal of Medicine* **335**, 1505–1513.

Holmes, D. T., Tariot, P. N. and Cox, C. (1998). Preliminary evidence of psychological distress among reservists in the Persian Gulf war. *Journal of Nervous Mental Disorders* **186**, 166–173.

Institute of Medicine (1995). *Health consequences of service during the Persian Gulf War: initial findings and recommendations for immediate action*. Washington, D.C.: National Academy Press.

Iowa Persian Gulf Study Group. (1997). Self-reported illness and health status among Gulf war veterans. *Journal of the American Medical Association* **277**, 238–245.

Kroenke, K., Koslowe, P. and Roy, M. (1998). Symptoms in 18 495 Persian Gulf war veterans: Latency of onset and lack of association with self-reported exposures. *Journal of Occupational and Environmental Medicine* **40**, 520–528.

Labatte, L. A. and Snow, M. P. (1992). Posttraumatic stress symptoms among soldiers exposed to combat in the Persian Gulf. *Hospital and Community Psychiatry* **43**, 831–833.

Landrigan, P. (1997). Illness in Gulf war veterans: causes and consequences. *Journal of the American Medical Association* **277**, 259–261.

Laor, N., Wolmer, L., Mayes, L. C., et al. (1996). Israel preschoolers under SCUD missile attacks. *Archives of General Psychiatry* **53**, 416–423.

Morgan, C. A., III, Hill, S., Fox, P., Isingham, P. and Southwick, S. M. (1999). Anniversary reactions in Gulf war veterans: A follow-up inquiry 6 years after the war. *American Journal of Psychiatry* **156**, 1075–1077.

Perconte, S. T., Wilson, A. T., Pontius, E., Dietrick, A. L. and Spiro, K. J. (1993). Psychological and war stress symptoms among deployed and non-deployed reservists following the Gulf war. *Military Medicine* **158**, 516–521.

Presidential Advisory Committee on Gulf War Veteran's Illnesses (1996). *Final report*. Washington, D.C.: U.S. Government Printing Office.

Robinson, S., Hemmendinger, J., Netanel, M., et al. (1994). Retraumatization of holocaust survivors during the Gulf War and SCUD missile attacks on Israel. *British Journal of Medical Psychology* **67**, 352–362.

Solomon, Z., Laor, N., Weiler, D., et al. (1993). The psychological impact of the Gulf war: a study of acute stress in Israel evacuees. *Archives of General Psychiatry* **50**, 320–321.

Southwick, S. M., Morgan, A., Nagy, L. M., et al. (1993). Trauma-related symptoms in veterans of Operation Desert Storm: a preliminary report. *American Journal of Psychiatry* **150**, 1524–1528.

Stretch, R. H., Marlowe, D. H., Wright, K. M., et al. (1996). Post-traumatic stress disorder symptoms among Gulf war veterans. *Military Medicine* **161**, 407–410.

Sutker, P. B., Uddo, M., Davis, M. J. and Ditta, S. R. (1995). War zone stress, personal resources, and PTSD in Persian Gulf war returnees. *Journal of Abnormal Psychology* **104**, 444–452.

Wolfe, J., Brown, P. J. and Kelley, J. M. (1993). Reassessing war stress: exposure and the Persian Gulf war. *Journal of Social Issues* **49**, 15–31.

Wolfe, J., Proctor, S. P., Davis, J. D., Borgos, M. S. and Friedman, M. J. (1998). Health symptoms reported by Persian Gulf war veterans 2 years after return. *American Journal of Industrial Medicine* **33**, 104–113.

Personality Processes

R J Larsen
Washington University, St. Louis, MO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R J Larsen, volume 3, pp 149–154, © 2000, Elsevier Inc.

Transactional Model of Stress
Summary

Glossary

Appraisal	The part of the transactional model related to an individual's cognitively assessing the demands of the event and determining his or her resources for coping with those demands.
Attributional style	The way an individual explains the causes of bad events.
Coping	The part of the transactional model related to strategies for enduring and overcoming stress.
Expectation	The part of optimism related to the anticipation of positive outcomes.
Internal/external style	The degree to which people blame the cause of bad events on themselves or on things outside of their control.
Neuroticism	A personality characteristic associated with such characteristics as emotional instability, physiological hyperreactivity, and a lowered threshold for the experience of a variety of negative emotions, such as anxiety, anger, irritation, depression, and distress.
Optimistic bias	The part of optimism related to an individual's perception that he or she is at a lower risk for negative events than the average person.
Optimistic explanatory style	The combination of having external, unstable, and specific explanatory styles with respect to bad events.
Primary appraisal	The estimate that an individual makes about the degree to which a particular event is a threat to his or her personal goals.
Secondary appraisal	The estimate that an individual makes about the degree to which he or she has the personal resources to cope with the demands of some particular event.
Self-disclosure	A coping strategy related to the expression of thoughts and feelings about upsetting events.
Self-efficacy	The part of optimism related to an individual's belief in his or her ability to do the behaviors necessary to achieve some desired outcome.
Specific/global style	The degree to which people blame the cause of bad events on factors that are particular to a given situation or on factors that are much more general and applicable to a large number of life situations.
Stable/unstable style	The degree to which people blame the cause of bad events on factors that are long lasting and relatively permanent or on factors that are more temporary and changeable.
Stress	The subjective feeling that is produced by events that an individual perceives as overwhelming and beyond his or her control.
Stress exposure	The part of the transactional model related to being in situations that have a high probability for evoking stress in the average person.
Stressor	An event capable of eliciting stress.
Transactional model of stress	A model that suggests that stress occurs when (1) an individual is exposed to a challenging event, (2) the person appraises the demands of the event and appraises his or her own resources for adjusting to those demands, and (3) the person initiates a strategy for coping.

Stress is the subjective feeling that is produced by events that individuals perceive as overwhelming and beyond their control. Events that typically elicit stress are called stressors. Not all people respond to stressors in the same way, however. Two people may undergo the same event, say a painful loss such as losing their job. One person may be devastated and completely overwhelmed, whereas the other may accept the event as a challenge and mobilize to take positive action. Differences between people in how they react to the same objective event highlight the fact that stress is not out there in the environment. Instead, stress really lies in the transaction between the individual and the characteristics of the environment. Personality processes may moderate this transaction; whether a particular event is experienced as stressful is determined, in part, by the personality characteristics of the person exposed to the event. This article reviews briefly the transactional model of stress and what is known about several stress-moderating personality processes.

Transactional Model of Stress

In general terms, the transactional model of stress, first formulated by psychologist Richard Lazarus,

holds that stress unfolds as follows: (1) an individual is exposed to a challenging event, (2) the person appraises the demands of the event (primary appraisal) and appraises his or her own resources for coping with those demands (secondary appraisal), and (3) the person initiates a strategy for coping. Outcomes then follow, such as a change in the situation or in the health, mood, or behavior of the individual. Personality processes can affect each of the three steps in this transactional model.

Stress Exposure and Neuroticism

Some personality variables are related to higher exposure levels to stress-initiating events. Many studies have documented the role of neuroticism in the exposure to events that elicit stress. Neuroticism is a broad and inclusive personality characteristic associated with such characteristics as emotional instability, physiological hyperreactivity, and a lowered threshold for the experience of a variety of negative emotions, such as anxiety, anger, irritation, depression, and distress. Individuals high in neuroticism (high-N) are frequently found to report greater rates of unpleasant events in their daily lives than low-N subjects, particularly social events such as family difficulties, problems with coworkers, and personal relationships that create conflict and distress. In general, high-N subjects report much more interpersonal stress in their lives than low-N individuals.

The process whereby high-N people come to be exposed to more stressful events is less well understood. One hypothesis is that high-N subjects evoke unfavorable reactions in others. That is, the general characteristics associated with neuroticism lead those around high-N people to react in ways that create stress for them. A second hypothesis is that high-N subjects simply elect to be in situations with higher conflict potential, seeking out those stress-provoking situations that are avoided successfully by low-N subjects. A third hypothesis is that high-N subjects have a lower threshold for reacting to potential stressors; that is, given the same event, high-N subjects exhibit a larger stress response than low-N subjects. Although a dozen or more studies report positive correlations between N and stress responses, more recent controlled laboratory studies using standardized stressors have found that high-N subjects react to identical stressors (e.g., an unpleasant mood induction) with more negative emotions than low-N subjects. For example, studies have shown that high-N people have more cortisol (a biological by-product of stress) in their systems following standardized stress tests than do low-N people. High-N subjects appear to react to challenging events in a way that amplifies their biological stress response. Other studies have documented that high-N subjects are more reactive to naturally occurring interpersonal and noninterpersonal events (e.g., financial stressors). Stress reactivity blends into the next step in the transactional model – how people cognitively process or appraise events can determine their degree of reactivity to that event.

Appraisal, Attributional Style, and Optimism

According to the transactional model of stress, in order for stress to be evoked for individuals, two cognitive events have to occur. In the first cognitive event, called primary appraisal, people perceive that the event is a threat to their personal goals; in the second necessary cognitive event, called secondary appraisal, people conclude that they do not have the resources to cope with the demands of the event. If either of these appraisals is absent – if people do not perceive the event as threatening or if they feel they have plenty of resources for coping with the threat – then stress is not evoked. If some event is perceived as being threatening to individuals' goals, yet people feels they have the resources necessary to meet that event, then they may experience the event more as a challenge than as stress. Alternatively, individuals may feel that they do not have the resources demanded by the event (secondary appraisal) and that the event is not important to their long-term goals (primary appraisal) and so may not respond with stress.

The appraisal step in the transitional model of stress highlights the importance of cognition in understanding how an event becomes stressful. Individuals may differ in how they cognitively process events such that some individuals experience more stress than others. Attributional style refers to individual differences in the way people explain the causes of bad events. For example, I might ask you why you just had a minor automobile accident. You might attribute your accident to your lack of driving ability, which is an internal attribution, due to you. Or you might attribute your accident to the car's being particularly difficult to drive because it is so unresponsive, which is an external attribution, due to something outside of your control.

Although people might explain any specific event with either internal or external attributions, most people tend to consistently lean in one direction or the other. That is, some people typically give internal explanations and others typically give external explanations. One way to think about attributional style is in terms of the question: Where does the person typically place the blame when things go wrong? For example, some people tend to blame themselves when something goes wrong (e.g., I crashed the car

because I am a clumsy driver) and others never blame themselves for any bad event (e.g., I crashed the car because the road was not built correctly).

In addition to the internal versus external style distinction, attributions can also be characterized as referring to stable versus unstable causes and to whether those causes are specific versus global. For example, say that you are asked why you got a low grade on an exam. An internal and stable attribution might be that you did poorly because you have low ability or are low in intelligence. This is a stable attribution cause because intelligence is thought of as a stable characteristic, as difficult to change. An internal but unstable attribution might be that you did poorly because you were tired that day. Being tired is an unstable attribution because you can get over that relatively easily. Global versus specific causes are whether the cause of the event is due to some global state of affairs or is specific. For example, in explaining a low grade in your statistics course, a global explanation may be that you are low in general intelligence; a more specific explanation may be that you just have trouble with numbers.

The three attributional style dimensions tend to be correlated in a particular pattern. Consequently, many psychologists prefer to view this pattern as a single personality dimension, which they call the optimistic versus pessimistic explanatory style. Optimists believe that they are in charge, that they are the master of their fates, and that their actions actually influence their outcomes in life. Pessimists, however, believe that they are fairly helpless when it comes to bad events and that nothing they can do will make much of a difference in preventing bad events. Consequently, optimists have the expectation that what they do will influence what happens to them, whereas pessimists expect that their behavior is not related to the outcome.

In addition to being related to an internal, unstable, and specific explanatory style, optimism is also related to other personality concepts. For example, optimism also includes the people's expectations that good events will be plentiful, and bad events rare, in the future; for example, optimistic people expect that they will achieve success in most areas of life.

Another related psychological concept, developed by psychologist Albert Bandura, is self-efficacy. This is the people's belief that they can do whatever is necessary to achieve some desired outcome. Self-efficacy also refers to the confidence people have in their ability to perform the actions needed to achieve some specific outcome. For example, if you lose the job you have had for many years, self-efficacy is your subjective belief and confidence that you can find a new job.

A third concept related to optimism concerns perceptions of risk. For example, imagine you are asked to estimate the probability of various unpleasant events happening to you using a scale from 0 to 100, where 0 means "it will never happen to me" and 100 means "it is certain to happen to me." The events you are asked to estimate include dying in a plane crash, being diagnosed with cancer, having a heart attack, and being hit by lightning. Optimists perceive that they are at lower risk for such negative events than the average person is. This is referred to as the optimistic bias, and it may actually lead people in general to ignore or minimize the risks inherent in life or to perceive the world as a fairly benign place.

Psychologists have theorized about the possible mechanisms whereby the optimistic explanatory style buffers stress. These mechanisms should not be viewed as competing hypotheses; instead, two or more of these mechanisms may be operating to buffer stress for optimists.

- Immune system. Researchers have shown that the immune systems of optimists respond to challenges better and with more strength than the immune systems of pessimists. Other researchers have examined how optimism relates to the progression of HIV. Although results are mixed, there are suggestions in the literature that optimists have a longer survival time after the development of AIDS symptoms, and that this may be due to stronger a immune function among optimists.
- Emotional mechanism. There is a very large literature showing that optimists are resistant to depression. It could be that optimism is related to stress buffering indirectly because it protects people from depression and the debilitating consequences of this emotional condition.
- Cognitive process. Optimism may be related to an extensive set of beliefs that people have about themselves and the world, and these beliefs may, in part, promote a more positive or adaptive perspective toward potential stressors. For example, researchers have shown that optimism is related to the belief that people can maintain and promote their own positive life circumstances as well as the belief that they can reduce various risks. Other studies have shown that optimists see the world as less filled with hassles and therefore feel less stress than pessimists.
- Social contact. Pessimists tend to be loners, and social isolation is a reliable predictor of general distress and dissatisfaction with life. Friends and family may provide the earliest medical feedback when things start going bad. For example, a friend might say, "You look stressed today," or a roommate might remark, "You are looking really tired these days; is

something wrong?" Because pessimists are isolated and avoid social contact, they may have less of this social feedback and hence not have this source of information about their health status.

- Direct behavior. Optimism may set in motion certain behaviors that lead directly to benefits for coping. The answer to how optimism and coping are linked may be as simple as that optimists act differently and take better care of themselves and their circumstances than pessimists do. Behavior may prove to be the most critical link in the connection between optimism and coping, simply because optimists do more of the right things (e.g., exercising, sleeping enough, and getting advice from others) and fewer of the wrong things (e.g., drinking and distracting themselves from problems).

As a caveat, most of the research on the association of optimism and stress is correlational, and therefore psychologists cannot pinpoint exactly which factors are responsible for the correlations; they can only speculate. Future research will help determine what are the most effective components of the optimistic personality style in producing better health and a longer life.

Coping and Self-Disclosure

What are some of the strategies that people use to overcome stress and the accompanying feelings of anxiety, being overwhelmed, and being afraid? One strategy studied by psychologists is self-disclosure, the tendency for people to express their thoughts and feelings to others.

Many theorists have suggested that keeping things to ourselves (not opening up to other people) may be a source of stress and ultimately may lead to psychological distress and physical disease. These theorists have further argued that being open with others about our feelings may promote coping. For example, talk therapy may work, in part, because through it patients uncover secrets and reveal what they have been keeping to themselves.

Psychologist James Pennebaker is a leader in the field of self-disclosure. In some early studies, he asked subjects to think of something upsetting or traumatic that had happened to them, something they had not discussed with anyone. He asked his subjects to write these secrets down. People wrote about embarrassing moments, sexual indiscretions, illegal or immoral behaviors, humiliations, and so on. It is interesting that all subjects came up fairly quickly with a secret that they had been keeping.

Not discussing some upsetting event can lead to problems because it takes physical energy to inhibit the thoughts and feelings associated with such events. It is not easy to keep a secret, and keeping something in, especially if it is a major trauma or an especially significant or upsetting event, takes a lot of energy. Over time the cost of keeping the secret builds and builds and can increase the likelihood of stress-related problems such as trouble sleeping, irritability, physical symptoms such as stomach aches and headaches, and even illness.

According to Pennebaker's early theorizing, telling the secret relieves the compounded stress of keeping it in. In other works, getting something off your chest is a relief from the cost of keeping it inside. Confronting the traumatic memory by telling someone or even writing about it releases the person from the work of keeping a secret. Letting the secret out by talking or writing about it lowers stress and improves the chances for coping and positive adjustment.

In his recent writings, Pennebaker has proposed a different mechanism of action for self-disclosure. By telling the story of some stressful event, people reinterpret the event, often finding a new way to think about it that lessens its impact. For example, a college student talking about a painful breakup may reinterpret the event to actually have some positive consequences (e.g., He really was not my type anyway). Telling others about misfortunes gives people an opportunity to reframe the event and to look for the silver lining in even the worst events.

Many researchers have conducted studies on the self-disclosure of stressful events or telling secrets. In some of these studies, researchers contacted subjects who had undergone a severe event, such as the loss of a spouse through accident or suicide. Subjects were asked how much they had discussed the event with friends, family, or other helping professionals, such as a priest or minister or therapist. Researchers also typically assessed the subjects' health during the time following the stressful event. Results typically showed that the more people had talked about the tragedy, the better their subsequent health. In other words, those who kept the trauma to themselves tended to suffer more problems subsequent to the event than those who disclosed their feelings to others.

In related research, experiments were conducted in which subjects were randomly assigned to conditions. In one condition, subjects were typically asked to recall and write about a distressing event. In the other condition, subjects were typically asked to write about some trivial topic, such as what they normally eat for breakfast. Results frequently showed that subjects writing about the traumatic event reported feeling more distress and discomfort while writing. However, follow-up surveys showed that subjects who wrote about a trauma typically had fewer illnesses compared to subjects who wrote

about trivial topics. Moreover, objective records (e.g., number of health-service visits) also suggest that subjects who self-disclosed about trauma did indeed have fewer illnesses than the subjects who wrote about trivial topics.

Other studies show that people who keep unpleasant information about themselves secret are more likely to develop anxiety or depression than those who tell someone. Often psychotherapists will ask their clients, especially those who have experienced a trauma or some extreme event, to talk or write about that trauma. Some psychologists even recommend that their patients keep a diary of the events in their life and how they are reacting to those events. Such a daily self-disclosure or private secret-telling helps people put their feelings into perspective and make some meaning out of the events in their life.

In summary, research on self-disclosure suggests that keeping traumatic events and the feelings about those events unexpressed can be stressful. Expressing emotions in words can, it appears, produce stress-reducing effects. Moreover, it appears that it does not matter how the feelings are put into words; it may not matter whether we talk to a trusted friend or relative, go to a caring psychotherapist, confess at our church, talk with our minister or rabbi, have a discussion with our husband or wife, or write it in a diary. Whatever form it takes, self-disclosure appears to be a personality process that can help people cope with stress.

Summary

The differences among people in how they respond to events exist because stress is not out there in the environment; instead, stress is in the subjective reaction of individuals to events. This is worth emphasizing because many people often refer to certain events as stressful, as if stress were an inherent characteristic of certain events. Instead, stress lies is in how individuals respond to events.

Personality processes may have their effects at one or more points in the stress response. This article presents the transactional model of the stress response, which emphasizes three steps: exposure, appraisal, and coping. Certain personality characteristics have been associated with each of these steps; for example, exposure has been associated with neuroticism, appraisal with attributional style and optimism, and coping with individual differences in self-disclosure. In reality, the steps of the stress response blend into one another, and so the personality processes presented here certainly overlap; for example, neuroticism may not only be associated with stress exposure, but may also be associated with coping efficacy. In addition, because of space constraints, there are other personality processes in the literature that were not discussed, such as the personality variables hardiness, locus of control, and emotional intelligence. Moreover, there are many individual differences related to health behavior (e.g., smoking, exercise, and social support) that may provide buffers for stress. Identifying and understanding individual differences in the stress response are fascinating and important goals for psychological research.

See Also the Following Articles

Coping Skills; Depression Models; Emotions: Structure and Adaptive Functions; Immune Response; Optimism, Pessimism, and Stress; Social Support.

Further Reading

Cooper, C. L. and Dewe, P. (2004). *Stress: a brief history.* Boston: Blackwell.

Dohrenwend, B. P. (1998). *Adversity, stress, and psychopathology.* New York: Oxford University Press.

Horowitz, M. (2001). *Stress response syndromes: personality styles and interventions.* Northvale, NJ: Jason Aronson.

Lazarus, R. S. (2006). *Stress and emotion: a new synthesis.* New York: Springer.

Lazarus, R. S. and Folkman, S. (1984). *Stress, appraisal and coping.* New York: Springer.

Moller, T. W. (1997). *Clinical disorders and stressful life events.* Madison, CT: International Universities Press.

Pennebaker, J. W. (1997). *Opening up: the healing powers of expressing emotion.* New York: Guilford.

Pennebaker, J. W. (2004). *Writing to heal: a guided journal for recovering from trauma and emotional upheaval.* Oakland, CA: New Harbinger.

Seligman, M. E. (2006). *Learned optimism: how to change your mind and your life.* New York: Vintage.

Pessimism *See:* Optimism, Pessimism, and Stress.

Phagocytes *See:* Macrophages; Glia or Neuroglia.

Pharmacological Treatments of Stress

O G Cameron
University of Michigan, Ann Arbor, MI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by O G Cameron, volume 3, pp 155–161, © 2000, Elsevier Inc.

Stress
General Clinical Psychopharmacology
Psychopharmacology of Stress
Conclusion

Glossary

Acute stress disorder (ASD)	The disordered human stress response that occurs in the first month after exposure to a trauma (stressor).
Fight or flight	Types of reactions in which organisms engage as adaptive responses to environmental demands.
General adaptation syndrome (GAS)	A name given to the organism's reaction to a stressor as described by Hans Selye. The GAS is comprised of three main features – thymic and lymphatic involution, gastric ulceration, adrenal hypertrophy.
Homeostasis	The hypothetical idealized state of the body (milieu interieur) in which all physiological functions are in their optimal basal or most unperturbed conditions.
Neurotransmitter	Any chemical in the brain or other part of the nervous system that carries information among nerve cells.
Posttraumatic stress disorder (PTSD)	The disordered human stress response that occurs subsequent to the first month after exposure to a trauma (stressor).
Psychopharmacology	The study of the pharmacological mechanisms of, or the therapeutic use of, drugs that affect the functions of the nervous system related to behavior and complex brain functions.
Stress	The complex of biological, psychological, and behavioral changes (that is, the stress response) that occurs as an attempt by an organism to respond adaptively (not necessarily successfully) to exposure to an environmental event (that is, the stressor); usually, but not invariably, has the connotation that the environmental event is detrimental or experienced as negative or unpleasant.

Stress

Stress is a widely used hypothetical concept that refers both to environmental events that evoke organismic responses and to the responses themselves. There is little systematic data on either the efficacy or appropriateness of using drugs to prevent or modify the normal stress response. In contrast, a persistent disorder of the stress response (posttraumatic stress disorder [PTSD]) can be at least partially relieved by medication. The most efficacious drugs are the antidepressants, especially the selective serotonin reuptake inhibitors (SSRIs). Furthermore, there are medications that can both relieve the acute stress response and diminish the likelihood of developing PTSD (benzodiazepines and β-adrenergic blockers).

The Definition of Stress

Before reviewing information concerning the psychopharmacological treatment of stress, it is essential to attempt to define stress. This is difficult, and not all individuals involved in stress research or the treatment of people with stress disorders are likely to agree with any particular definition. It is important, nonetheless, to identify basic issues about the definition of this concept as it has been used over the past half century.

Stress has generally been defined in one of two fundamental ways. In the first, stress is an environmental event or circumstance that produces an effect of a particular kind on the organism. That effect has been called the stress response (see **Evolutionary Origins and Functions of the Stress Response, Genetic Polymorphisms in Stress Response, Seasonal Changes in Stress Responses, Sex Differences in Human Stress Response**).

Alternatively, stress has been used to refer to the effects that occur in the organism. In this case, the

environmental event is typically called the stressor. Whenever stress is discussed, one must know in which way the speaker or writer is using the term. For the psychopharmacological treatment of stress, the word is typically used in the second sense, that is, to modify the effect of a stressor on the organism. If one wished to treat stress in the other sense, that would involve not changing the organism's response with treatment, but rather changing the environment itself. In other words, it would involve developing strategies to diminish the occurrence of stressors (a difficult thing to do in real life, and unlikely to involve drugs).

Circularity in the Definition of Stress

An ambiguity or circularity exists in the definitions of stress. In a situation in which an environmental event is quite different from normal circumstances (in other words, unambiguously traumatic), and especially if the event is not avoidable or escapable by the organism (assuming it is a negative event or circumstance, which is usually – but not always – considered to be the case in stress), most people would agree that a stress or stressor has occurred. These are the kinds of events often used in stress research, especially with animals. However, in more typical real-life circumstances, it is frequently not as clear whether a stressor has occurred. For example, while some people find sky diving exhilarating and enjoyable, others find it very frightening. Is it a stressor? Is it a stressor only for those who report a negative subjective response to it? Is it a stressor for anyone who reports a strong subjective reaction, positive or negative? Is it a stressor for everyone, but some people cope with it more effectively? How can stress be unambiguously defined and recognized?

In an attempt to avoid this dilemma, some have advocated using the organism's response instead of the environmental event as the defining characteristic. Classic markers have included physiological reactions such as changes in heart rate, blood pressure, and sudomotor response (sweating), as well as changes in hormones, especially glucocorticoids (such as cortisol) and epinephrine (adrenaline). In this situation, there is a risk that the physiological (or other) changes will be equated with stress rather than being a marker of it. Unless an external event or circumstance is so extraordinary that all observers agree to its divergence from normal circumstances, this definitional problem appears to be almost inevitable. The issue of normal versus abnormal stress is important in understanding the treatment of stress, as described next.

Stress Normal and Abnormal, and Homeostasis

In the middle of the nineteenth century Claude Bernard defined the concept of the milieu interieur, his point being that the organism regulates all of the physiological processes going on in the body (under the skin). Early in the twentieth century, Walter Cannon hypothesized that the organism seeks homeostasis, that is, there is an optimal status for the milieu interieur, and the organism attempts to maintain that set of conditions. While there is much truth to this concept, it has been found to be an oversimplification. For example, most physiological processes have been shown to fluctuate normally over predetermined time cycles, for example, the circadian (around 24 h) rhythm and the monthly menstrual cycle.

In the context of homeostasis, ignoring the definitional problem, stress can be understood as a perturbation of this set of psychological, physiological, biochemical, and even gene expression states. A fundamental question thus arises: is the stress response – the organism's response to the stressor – a normal reaction? And, if so, what is its function?

Cannon gave an answer to this question, although he was not thinking exactly in terms of the stress response. He viewed the various physiological changes that occur as preparing the organism for fight or flight, in other words, for action. Extensive study of the multitude of reactions that occur during the response to the stressor has generally been consistent with this hypothesis. In short, Cannon was largely correct.

Perhaps the scientist most closely associated with the concept of stress was Hans Selye. He called it the general adaptation syndrome (GAS). He viewed the stress response somewhat differently, not clearly as an adaptive mechanism but as evidence of dysfunction, at least in its more extreme form. In short (admittedly an oversimplification), both Cannon and Selye viewed the changes that are now called the stress response as normative, that is, always or almost always occurring when a stressor is present. However, Cannon viewed the response as adaptive and healthy, while Selye viewed it, at least in its more severe form, as destructive.

Disorders of Stress

Whether or not one views the stress response as adaptive, it does appear that it is normal; in other words, it does always or almost always occur when a stressor is present, at least if the stressor is intense enough. It is also true, however, that a minority of individuals experience stress responses that are abnormal in character, intensity, and/or duration.

Clinical descriptions of detrimental psychological changes occurring in response to traumatic events go back at least as far as the Civil War, and probably much earlier. Phenomena previously identified by such terms as soldier's heart and shell shock refer to such changes occurring in association with traumatic events in war, but these negative psychological reactions can occur to a wide variety of severely traumatic events. The current psychiatric diagnosis nomenclature, as sanctioned by the American Psychiatric Association, is called the *Diagnostic and Statistical Manual of Mental Disorders,* fourth edition (DSM-IV). In it, these detrimental changes are defined as disorders, and are classified as anxiety disorders. The two specifically defined disorders are acute stress disorder (ASD) and PTSD. The diagnostic criteria for these two disorders are listed in Table 1.

The major differences between the two types of stress disorders are (1) the duration criterion, (2) the dissociation criterion, which is present only in ASD, and (3) the exclusion criterion for substance use, etc., which also is only in ASD. However, for the purpose of this article, the main issue is the fact that both stress disorders are just that, disorders. In other words, these are not descriptions of how the normal or modal person responds to a traumatic event. These are descriptions of how responses occur when they go wrong in some way, definitions of abnormal responses. Research has shown that the majority of people who experience most kinds of traumatic events do not develop stress disorders. Predictors of who will develop these syndromes after experiencing a trauma are beginning to be identified. Treatment is usually recommended only for the individuals who develop a stress disorder; the question of whether or not (and under what circumstances) to treat the stress response – that is, to treat any individual exposed to a severe stressor – is revisited at the end of this article.

Table 1 DSM-IV diagnostic criteria for acute stress disorder (ASD) and posttraumatic stress disorder (PTSD)

Acute Stress Disorder (ASD)
A. Exposure to a traumatic event
B. Three or more of five dissociative symptoms
C. Persistent re-experiencing of the event
D. Marked avoidance of stimuli that arouse recollections
E. Marked symptoms of anxiety or arousal
F. Clinically significant distress or impairment
G. Occurs within 4 weeks of the event and lasts between 2 and 28 days
H. Not due to a substance, medical condition, or other psychiatric disorder
Posttraumatic Stress Disorder (PTSD)
A. Exposure to a traumatic event
B. Persistent re-experiencing of the event in one or more of five ways
C. Persistent avoidance and/or numbing in three or more of seven ways
D. Persistent increased arousal in two or more of five ways
E. Duration is more than 1 month
F. Clinically significant distress or impairment

General Clinical Psychopharmacology

Before discussing the treatment of stress disorders with psychopharmacological agents, a brief introduction to psychotropic drugs is in order. Major classes of these drugs are listed in Table 2. The two classes that are most useful in the treatment of PTSD – antidepressants and anti-anxiety drugs – are listed separately. The remaining drugs are listed only as the major classes rather than the individual drug groups.

In many branches of medicine, drugs are listed in classes related to the first or most well defined use for which they have been shown to be effective. For example, a number of drugs with various chemical structures – heterocyclics, monoamine oxidase inhibitors (MAOIs), SSRIs – are all called antidepressants because (1) they all demonstrate an antidepressant effect, (2) this is a very important effect, and (3) this is often the effect for which each drug was first demonstrated to have benefit. Despite this, drugs in the group of antidepressants are beneficial for a variety of disorders. For example, the heterocyclics are also used to treat enuresis (bed wetting), some kinds of pain, obsessive-compulsive disorder, and panic anxiety disorder. MAOIs are useful, for example, for panic disorder and social phobia. SSRIs treat a wide

Table 2 Classes and examples of psychotropic drugs

Antidepressants
1. Heterocyclic
2. Monoamine oxidase inhibitors (MAOIs)
3. Selective serotonin reuptake inhibitors (SSRIs)
4. Other agents
Anxi-anxiety drugs
1. Benzodiazepines
2. Other
Other classes
1. Mood stabilizers
2. β-adrenergic antagonists
3. α2-adrenergic agonists
4. Antipsychotics
5. Stimulants
6. Cognition enhancers
7. Narcotic analgesics
8. Drugs for substance abuse prevention

variety of disorders such as panic disorder and the eating disorder bulimia nervosa. The antidepressant drugs, especially the SSRIs, are the drugs that have been demonstrated most clearly to have benefit for people with stress disorders.

The other class of drugs for which there is controlled-trial research concerning treatment of the stress disorders is the benzodiazepines. These drugs are used most widely as anti-anxiety agents. They are also used as sedatives, muscles relaxants, hypnotics (sleeping pills), and anticonvulsants (anti-seizure). This class includes such drugs as alprazolam, diazepam, and clonazepam.

As can be seen in Table 2, there are several other groups of drugs classified as psychopharmacological agents. Either these other groups have not been tested for the treatment of stress disorders, or the studies have not been controlled trials. Table 2 does not contain all drugs or drug groups with psychotropic effects that are potentially useful, however; undoubtedly new drugs will be added in the future.

Psychopharmacology of Stress

In addressing the question of how to treat stress with medications (or otherwise), the first issue that arises is which meaning of stress is under consideration. In addition to PTSD and ASD, many psychiatric syndromes such as schizophrenia, depression, and various anxiety disorders, as well as various medical disorders, are often considered as involving stress. The concept of stress is used very broadly, while research specifically related to DSM-IV defined disorders of stress is much more restricted. Only a small number of controlled studies have been published.

Acute Stress Disorder (ASD)

There has been less research on the treatment of ASD than on that of PTSD, but two issues have emerged. First, medications with rapid onset of effect that can reduce anxiety and/or arousal-like symptoms appear to be appropriate sometimes. Second, a number of acute risk factors for development of PTSD have been identified, including the occurrence of acute stress symptoms and especially adrenergic arousal-like symptoms (mainly, elevated heart rate soon after the stressor, and also elevated urinary epinephrine). (History of childhood trauma, severity and persistence of physical injury, and abnormalities of cortisol secretion are among the other implicated predictors.) Specific medications potentially useful to treat acute stress symptoms include the benzodiazepines and the β-adrenergic blocking drugs (which also can reduce the likelihood of developing PTSD).

Posttraumatic Stress Disorder (PTSD)

Even though the relationship between the experience of a highly disruptive event and the development of post-event PTSD symptoms has been known for many years, systematic research into the use of psychopharmacological agents to treat this type of symptoms has been ongoing for only about 20 years. The number of studies is relatively small, and the number of controlled studies is smaller still. One reviewer has summarized these studies by concluding that drug treatment has modest but meaningful effects.

Specific drugs for PTSD Controlled studies of drug treatment of PTSD have mainly evaluated drugs in the antidepressant class. Among the heterocyclic antidepressants, there is evidence for benefit from imipramine, amitriptyline, and probably chlorimipramine. It appeared that approximately two-thirds of the treated individuals experienced meaningful improvement. The symptoms most likely to improve were intrusive symptoms such as distressing recollections or dreams. Avoidance symptoms were typically less responsive. On the other hand, desipramine did not appear to be as helpful. It has been observed that the drugs that were most helpful have a more prominent effect on the neurotransmitter serotonin in the brain, while less beneficial drugs have a more prominent effect on the brain neurotransmitter norepinephrine.

Antidepressants from two other groups, MAOIs and SSRIs, have also been analyzed in controlled studies. Results with MAOIs have been mixed, but have also supported their usefulness for intrusive symptoms. SSRI drugs have shown robust benefit, not only for the intrusive symptoms but also for avoidance and numbing. For all of these medications, early improvement was associated with a better ultimate outcome, but significant benefit nonetheless typically required 5–8 weeks or more, and further improvement was sometimes observed well beyond 8 weeks.

The benzodiazepine alprazolam was assessed in a controlled trial. It was less beneficial than the antidepressants. Hyperarousal symptoms did, however, show some improvement. Other benzodiazepines are also likely to show benefit for the hyperarousal symptoms.

In addition to the controlled studies, a number of uncontrolled treatment trials have been reported, and a number of different drugs and drug classes have been assessed. Although these results must be viewed as preliminary, promising drugs include the anticonvulsant mood stabilizers carbamazepine and valproic acid. Other drugs also probably have a role as second-tier agents to be used either in combination

with the first-tier drugs or when the first-tier drugs fail. These include the mood stabilizer lithium, the β-adrenergic antagonist drugs, the α2 agonists, the nonbenzodiazepine anti-anxiety agent buspirone, the sedating antidepressant trazodone, and opiate antagonists (possibly beneficial for numbing symptoms). Antipsychotic drugs are not indicated for PTSD unless psychotic symptoms are present.

Issues and factors in drug treatment of PTSD The definition of PTSD includes three groupings of symptom types: (1) intrusive, (2) avoidance and numbing, and (3) hyperarousal. The heterocyclic antidepressant drugs (primarily those that predominantly affect serotonin) and the MAOIs have been shown to be most beneficial for the intrusive symptoms. SSRIs are also beneficial for the intrusive symptoms, as appears to be the mood stabilizer anticonvulsant carbamazepine. Considering all three types of symptoms, SSRIs appear to be the most beneficial. In general, the intrusive symptoms are probably the ones most responsive to drug treatment.

The avoidance and numbing symptoms appear to be the most difficult to control. They are least likely to improve, and they require the longest treatment – many weeks – before improvement occurs, if it occurs at all. The antidepressant drugs, mainly the SSRIs, are most likely to produce improvement for these symptoms.

Hyperarousal symptoms seem intermediate in their likelihood of response. Benzodiazepines, buspirone, and the β-adrenergic antagonists are most specifically efficacious for the hyperarousal symptoms.

Several other issues in the drug treatment of PTSD are noteworthy. Not only do these drugs take longer to provide benefit for PTSD than they usually do for depression (2 months or more versus approximately 1 month), but they often require higher doses. Often only partial symptom relief is attained with the drugs now available. A total treatment program, including nonpharmacological treatments, is almost always appropriate. The presence of comorbid disorders such as depression, other anxiety disorders, and personality disorders modulates treatment outcome, although the direction of effect is variable. For example, some research suggests that the presence of depression improves prognosis, while other research suggests prognosis is worse when depression is present, possibly because of the increased overall severity due to two co-existing (comorbid) disorders. (Medication benefit does not depend on the presence of a comorbid disorder.) The presence of a personality disorder, especially antisocial personality, makes prognosis worse. The source of the trauma (for example, combat versus non-combat related) can influence outcome. Finally, it is likely that the earlier that treatment is provided after the PTSD syndrome develops (or even in the first month after the trauma – ASD), the more improvement can be anticipated.

Table 3 Factors and goals of drug treatment of PTSD[a]

1. Improve intrusive symptoms
2. Control avoidance and numbing symptoms
3. Reduce hyperarousal symptoms
4. Provide distance between the individual and intensity of the emotional experience
5. Relieve depressive symptoms
6. Minimize irritability and aggressiveness
7. Gain control of dissociative and psychotic symptoms if they occur
8. Facilitate efficacy of psychotherapy

[a]Modified from Vargas and Davidson (1993).

It has been suggested that the phasic (intermittent) PTSD symptoms are more responsive to treatment than the tonic (persistent) symptoms. A list of the ways in which drug treatment can be helpful in the management of PTSD appears in Table 3.

Stressful Circumstances: Should Normal Stress Be Treated?

The research reviewed previously demonstrates that medications are beneficial, but not usually curative, for PTSD. Older research had suggested the use of sedative drugs for abreaction for ASD, a treatment now not generally recommended. Some clinicians have suggested that sedation, possibly including sedative effects from benzodiazepines, might be detrimental, although the anti-arousal, anxiolytic effects can be acutely beneficial. These data lead to the question, can the normal stress response itself be diminished or prevented effectively (with pharmacological and/or other means)? Should it be?

If one views the stress response as Cannon conceived of it (that is, the fight-or-flight reaction), it would seem that prevention of this normal physiological and behavioral response to environmental stressors would be detrimental rather than beneficial. On the other hand, if one views the stress response as Selye did, that is, as a reaction that at least in its more severe form is harmful to the organism, preventing it would be appropriate. As pointed out by Yehuda and McFarlane, there is considerable confusion about the relationships among stressors, potential detrimental effects of the stress response itself, and the development of PTSD. For example, although it is widely believed that stress has many negative consequences on health, the relationship between stress and the development, maintenance, and/or

exacerbation of specific medical symptoms and syndromes remains at best only partially understood. Provocative preliminary hypotheses exist, such as the apparent detrimental effects of chronic elevations of corticosteroid stress hormone substances on specific areas of the brain (especially the hippocampus, a structure very important for normal memory function), but much more needs to be known to clarify these linkages. It seems that the answer to the question of whether to treat normal stress must await not only additional data, but also improved clarity concerning the definition(s) and use of the word and concept of stress as applied to humans in clinical situations. What is clear is that abnormal stress responses call for intervention.

While investigators and clinicians struggle to obtain greater conceptual clarity in the field of stress research and treatment, new data have begun to address the question of early pharmacological intervention for abnormal stress responses. Some early risk factors, predictive of increased likelihood of the development of PTSD, can now be used to lower that risk. Specifically, the presence of an elevated heart rate or other evidence of significant adrenergic activation soon after the trauma appears to call for measures, pharmacological and/or nonpharmacological, to promptly lessen this activation.

Conclusion

Disordered chronic responses to significant stressors (PTSD) are treatable with medications, although most people with the disorder do not experience a full remission from drug treatment, and some symptoms (avoidance and numbing) are harder to treat than others (intrusive, hyperarousal). Most individuals need a treatment program including pharmacological and nonpharmacological methods to derive maximal benefit. Prophylactic drug treatment given soon after trauma exposure to reduce arousal can diminish the risk of subsequent development of PTSD. The larger question of whether or not stress per se is treatable (that is, the normal stress response is diminished or prevented), and whether or not that would be advantageous or disadvantageous to the organism, depends in part on which definition of stress is under consideration.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Alarm Phase and General Adaptation Syndrome; Evolutionary Origins and Functions of the Stress Response; Fight-or-Flight Response; Genetic Polymorphisms in Stress Response; Homeostasis; Seasonal Changes in Stress Responses; Sex Differences in Human Stress Response; Selye, Hans; Stress, Definitions and Concepts of; Sympathetic Nervous System.

Further Reading

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.) (DSM-IV). Washington, D.C.: American Psychiatric Press.

Brady, L. S. (1994). Stress, antidepressant drugs, and the locus coeruleus. *Brain Research Bulletin* **35**, 545–556.

Cannon, W. B. (1939). *The wisdom of the body*. Philadelphia: W.W. Norton.

Davidson, J. (1992). Drug therapy of post-traumatic stress disorder. *British Journal of Psychiatry* **160**, 309–314.

Goldstein, D. S. (1994). Stress and science. In: Cameron, O. G. (ed.) *Adrenergic dysfunction and psychobiology*, pp. 179–236. Washington, D.C.: American Psychiatric Press.

Katz, L., Fleischer, W., Kjernisted, K., et al. (1996). A review of the psychobiology and pharmacotherapy of posttraumatic stress disorder. *Canadian Journal of Psychiatry* **41**, 233–238.

Marmar, C. R., Neylan, T. C. and Schoenfeld, F. B. (2002). New directions in the pharmacotherapy of posttraumatic stress disorder. *Psychiatric Quarterly* **73**, 259–270.

Pitman, R. K., Sanders, K. M., Zusman, R. M., et al. (2002). Pilot study of secondary prevention of posttraumatic stress disorder with propranolol. *Biological Psychiatry* **51**, 189–192.

Selye, H. (1956). *The stress of life*. New York: McGraw Hill.

Tomb, D. A. (ed.) (1994). *Psychiatric clinics of North America, vol. 17: Post-traumatic stress disorder*. Philadelphia: W.B. Saunders Company.

Turnbull, G. J. (1998). A review of post-traumatic stress disorder, part II: treatment. *Injury* **29**, 169–175.

Vargas, M. A. and Davidson, J. (1993). Post-traumatic stress disorder. In: Dunner, D. L. (ed.) *Psychiatric clinics of North America, vol. 16: psychopharmacology II*. Philadelphia: W.B. Saunders Company.

Weiner, H. (1992). *Perturbing the organism*. Chicago: University of Chicago Press.

Yehuda, R. (ed.) (2002). *Psychiatric clinics of North America, vol. 25: Recent advances in the study of biological alterations in post-traumatic stress disorder*. Philadelphia: W.B. Saunders Company.

Yehuda, R. and McFarlane, A. (1995). Conflict between current knowledge about posttraumatic stress disorder and its original conceptual basis. *American Journal of Psychiatry* **152**, 1705–1713.

Pheromones

A Kumar, C A Dudley, S Chakravarty and R L Moss
University of Texas Southwestern Medical Center, Dallas, TX, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C A Dudley, A Kumar and R L Moss, volume 3, pp 162–168, © 2000, Elsevier Inc.

Chemical Communication and Social Odors
The Vomeronasal Organ
Vomeronasal Organ Response to Olfactory Stimuli and Olfactory System Response to Pheromonal Stimuli
Stress and Reproduction
Stress Signaling Mediated by Pheromones

Glossary

Accessory olfactory bulb (AOB)	A neural structure located at the dorsocaudal end of the main olfactory bulb but distinguished from the main olfactory bulb by the types of signals detected and by its central nervous system projection sites.
c-Fos immunocytochemistry	A method for measuring cellular activation by detecting the expression of the c-fos oncogene.
Chemoarchitecture	The relationship between cellular topography and the cellular machinery used in signal detection and processing.
Chemosensory stimulation	An environmental stimulus detected by a sensory system via ligand–receptor binding.
Gonadotropin-releasing hormone (GnRH)	A decapeptide produced in the hypothalamus, responsible for stimulating the release of luteinizing hormone from the pituitary gland; also referred to as LHRH.
Sociogenic stress	A set of stresses caused by conspecific interactions.
Vomeronasal organ (VNO)	a specialized chemosensory structure for pheromone signaling.

Chemical Communication and Social Odors

Interorganism communication is an indispensable feature of species survival, and the signals employed for this purpose are quite diverse. Among the various senses employed, chemosensation (the detection and transduction of chemical environmental stimuli) is perhaps the most primitive. In addition to aiding in locating nutrition, chemosensory transduction allows an animal to find suitable mating partners and avoid dangerous predators. Pheromones are compounds used for intraspecies chemical communication that are secreted or excreted by one individual and elicit behavioral or physiological responses in the recipient. However, the great diversity of signals used in chemical communication by vertebrates suggests that this communication is not mediated exclusively by pheromones. A multiplicity of signals may be present in the same secretion, and it is unlikely that the response to this complex signal is mediated exclusively by a single sensory system. Johnston has recently suggested the term chemical signals for the wide array of odors regulating vertebrate behavior and physiology. Indeed, various chemosensory structures have been found in vertebrates that transduce chemosensory cues. Pheromones act on the recipient through olfaction, ingestion, or absorption; however, all mammalian pheromones studied so far act through olfactory pathways. Olfactory-mediated pheromonal communication in mammals is used to organize a wide range of social behaviors, such as mating, dominance, aggression, and maternal bonding, and hence these cues are also frequently referred to as social odors.

Mammalian pheromones elicit both long-lasting effects that bring about changes in the endocrine state of the recipient animal and short-term rapid effects on the recipient's behavior. For example, male mouse pheromones have been shown to induce advanced onset puberty in female counterparts (the Vandenbergh effect), the onset of estrus (the Whitten effect), and the termination of pregnancy (the Bruce effect) by an increase in luteinizing hormone (LH) and a decrease in prolactin. Female mouse pheromones cause a delay in the onset of puberty and a suppression of estrus in other female mice by increased release of prolactin. These delayed neuroendocrine effects on reproductive physiology and behavior have been termed the primer effect of mammalian pheromones.

The faster-acting releaser effects are elicited immediately, and these pheromones have been implicated in the regulation of intermale aggression, initiation of male ultrasonic vocalization and copulatory behavior, induction of lordosis (the receptive posture adopted by female rodents anticipating sexual intercourse), and parental care. Releaser effects are also

modulated by the hormonal status of the animal, which is itself under pheromonal regulation. Gonadotropin-releasing hormone (GnRH) and LH play crucial roles in regulating these pheromone-induced fast behavioral responses.

The Vomeronasal Organ

Pheromone molecules are chemically unrelated and are contained in body fluids such as urine, sweat, and sebum. Pheromones are detected by peripheral sensory neurons sequestered in a tiny cigar-shaped bony capsule called the vomeronasal organ (VNO) or Jacobson's organ (Figure 1a). Located bilaterally on the posterodorsal aspect of the nasal septum, this tubular structure opens into the nasal or oral cavity via the vomeronasal duct, allowing chemical stimuli to reach VNO neurons from both the nose and mouth. In some mammals such as ungulates and felines, the males display a behavior called flehmen, which consists of curling of the upper lip and throwing the head up after investigating the female's genital area. This helps in facilitating the transport of the odorous compounds in female urine into the VNO of male, thus allowing the detection of the female's receptive state. In rodents, however, a pumping activity of the VNO associated with the autonomic afferent system appears to actively draw pheromone compounds into the lumen of the VNO, where they bind to receptors. These receptors lie on microvillous extensions of the dendrites of the vomeronasal receptor neurons (VNRNs). Patch clamp recordings from dissociated VNRNs reveal that they possess voltage-gated inward and outward currents consistent with the ability to generate action potentials. VNRNs from the mouse respond to putative pheromones from mouse urine with the generation of an outward current accompanied by a decrease in membrane conductance (Figure 2). In a current clamp mode, pheromone application causes hyperpolarization of the membrane. The binding of pheromone molecules to their specific membrane receptors activates an intracellular signaling cascade involving G-proteins ($G_{\alpha i2}$ in the apical neurons, $G_{\alpha o}$ in basal neurons), phospholipase C-β (PLC), and diacylglycerol (DAG). DAG, the second messenger, is the gating factor for a cation channel known as transient receptor potential channel 2 (TRPC2), whose opening leads to an influx of Na^+/Ca^{2+} that causes membrane depolarization. The spread of membrane depolarization along the dendrite leads to the activation of voltage-gated Na^+ and K^+ channels, with the consequent generation of an action potential. The propagation of impulse along the nerve fibers allows VNO neurons to communicate with the first central processing unit, the accessory olfactory bulb (AOB).

Much insight into the molecular and anatomical components of VNO-signal transduction has been gained from the study of mutant mice. Mice that lack a functional TRPC2 channel show normal olfactory function but fail to respond to urinary vomeronasal stimuli. These mice do not display male–male

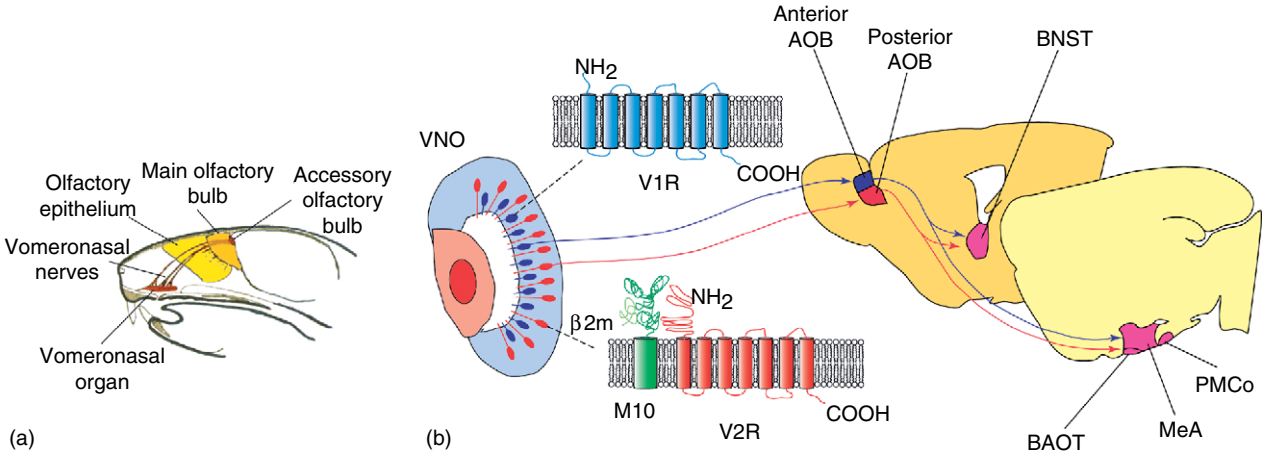


Figure 1 Rodent vomeronasal organ (VNO). a, Sagittal view showing the location in relation to the main olfactory epithelium; b, anatomy and central neural projections, with the sensory vomeronasal receptor neurons (VNRNs) shown in a coronal cross section of the VNO. Two discrete classes of VNRNs (V1Rs and V2Rs), G-protein-coupled receptors, project separately to the anterior and posterior accessory olfactory bulb (AOB), respectively. The V2Rs differ markedly from the V1Rs in that they have a large extracellular N-terminus and form a receptor complex with members of the M10 family of major histocompatibility complex (MHC) class Ib molecules and β_2-microglobulin (β2m). The information from the V1Rs and V2Rs are processed in the AOB and then converges in overlapping projections to the bed nucleus of the stria terminalis (BNST), bed nucleus of the accessory olfactory tract (BAOT), medial amygdala (MeA) and the posteromedial cortical nucleus (PMCo) of the amygdala. COOH, From Brennan, P. A. and Keverne, E. B. (2004), Something in the air?: new insights into mammalian pheromones, *Current Biology* **14**, R81–R89, with permission.

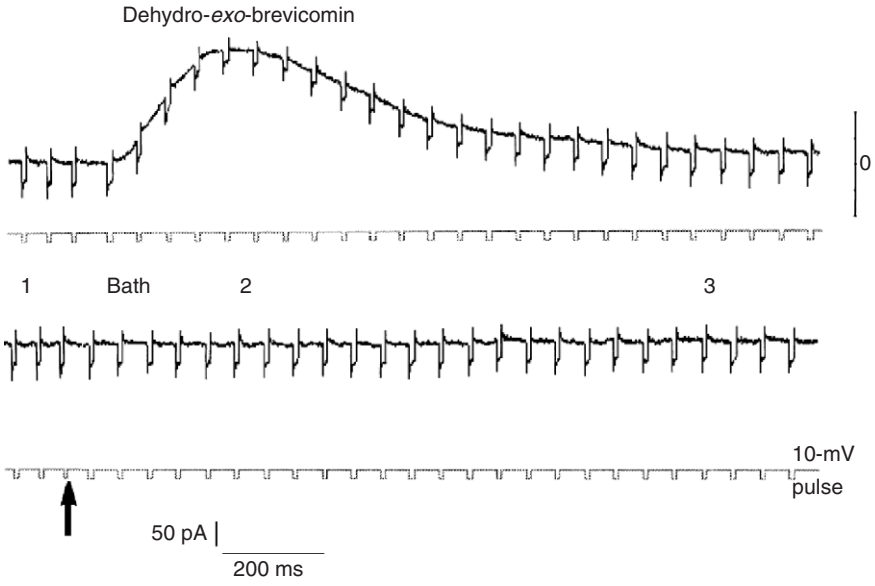


Figure 2 Application of the putative pheromone dehydro-*exo*-brevicomin (DHB) to a VNRN evoked an outward current accompanied by a decrease in conductance. Depolarizing pulses of 10 mV were applied continuously before (1), during (2), and after (3) DHB or bath. In the presence of DHB, the amplitude of the current pulses was decreased. From Moss, R. L., Flynn, R. E., Shen, X. M., et al. (1997), Urine-derived compound evokes membrane responses in mouse vomeronasal receptor neurons, *Journal of Neurophysiology* **77**, 2856–2862, with permission.

aggression and engage in sexual and courtship behaviors toward both males and females. They also have deficits in maternal aggression toward male intruders. This lack of aggressiveness in these mutant mice is in accordance with similar findings in mice following VNO lesions. The mouse VNO has two broad receptor gene families, V1Rs and V2Rs, comprising over 250 receptors genes. When the genes for 16 of the V1R family clustered together were deleted using chromosomal engineering (which amounted to 12% of the V1R population), male mice were impaired in their sexual behavior to females, with fewer males initiating sexual activity toward females than the genetically intact controls. These mice, too, fail to initiate maternal aggression toward intruder males, as was found in TrpC2-deficient mice, although there were no differences in the aggressive behavior of these V1R cluster-deficient males. These and other studies have yielded important insights into the functioning of the vomeronasal system and have helped researchers appreciate the complexities of this chemosensory system.

Chemoarchitecture of the Vomeronasal Organ to Accessory Olfactory Bulb Projection

The diversity in electrical responses of VN receptor neurons to putative pheromones, as well as the vast array of behaviors subjected to pheromone regulation, may be related to the fact that VNRNs appear to be divided into at least two chemoarchitecturally distinct populations: VNRNs located in layers close to the lumen of the VNO express $G_{\alpha i2}$ and project to the anterior portion of the AOB; VNRNs located more basally express $G_{\alpha o}$ and project to the posterior part of the AOB. The significance of this division has become more apparent with studies identifying two families of VNO receptor genes, with V1Rs and V2Rs expressed in $G_{\alpha i2}$ and $G_{\alpha o}$ VNRNs, respectively. This expression pattern suggests that there exist two functionally distinct processing units responsible for relaying signals from the VNO to the AOB.

Until now, the physiological basis for chemoarchitecturally segregated populations of VNO neurons was largely unexplored. Presuming that both populations of VNO cells are involved in pheromone detection, they may recognize separate categories of pheromones (volatile vs. nonvolatile, or primer vs. releasing) or may be involved in the mediation of separate behavioral responses. In the male mouse, the VNO is involved in at least two distinct behaviors: reproduction and aggression. Experiments have demonstrated that neurons in the anterior portion of the AOB of the male mouse are preferentially activated by exposure to female-soiled bedding, as opposed to familiar male-soiled bedding. The differential activation was dependent on an intact VNO and was not observed in the main olfactory bulb (MOB). These data indicate that the position of a vomeronasal receptor cell within the VNO is reflected at the level of the AOB by a functional division along the anterior–posterior axis.

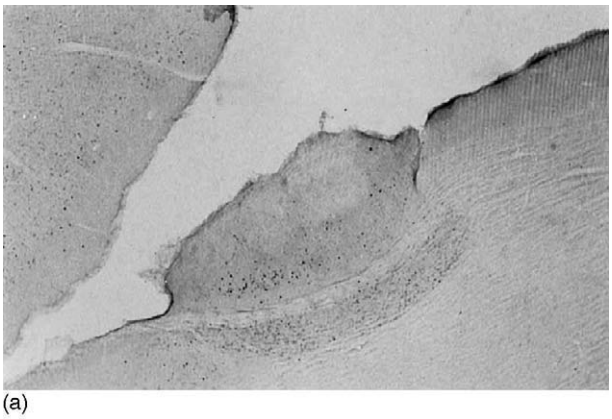
Recently, Katz's group performed the first recordings of pheromonal responses in the AOB of awake-behaving mice that demonstrated that responses of the AOB mitral cells are highly selective to strain and the sex of conspecifics. It is also possible that nuclei downstream of the AOB, such as the medial amygdale (mAMG), may also be preferentially activated. The nature of the information processing by V1Rs and V2Rs is yet to be understood, but the anterior and posterior AOB projections completely overlap at the level of amygdala, the bed nucleus of the accessory olfactory tract (nAOT), and the bed nucleus of the stria terminalis (BNST).

Central Neural Pathways of the Vomeronasal System

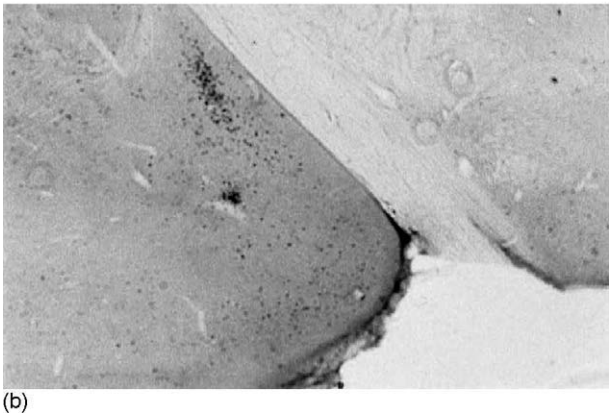
The anatomical pathway linking pheromone detection in the VNO to the hypothalamic centers controlling sexual and reproductive events has been well defined by studies using transsynaptic tracers. VNRNs project directly to the AOB, where they synapse onto mitral cell dendrites. The mitral cell axons project to the mAMYG, BNST, and nAOT. Connections between the mAMYG and the preoptic area (POA) and between the mAMYG and the ventromedial hypothalamus (VMH) complete the circuit by which pheromones may modulate the hypothalamic and/or limbic centers involved in reproductive behavior. Lesions of the VNO, AOB, mAMYG, or VMH significantly impair the ability of the female rat to display increases in sexual receptivity in response to male sexual contacts. The possibility of direct projections from the VNO to the POA and VMH has received some support and may provide an additional pathway for pheromonal control over GnRH-secreting neurons located in the hypothalamus. The surgical removal of the VNO adversely affects the copulatory behavior in males and lordosis behavior in female rodents. GnRH has been shown to facilitate mating behavior in the female rat, and GnRH has also been demonstrated to restore mating behavior in the female rat and male hamster after VNO removal.

Activation of Vomeronasal Neural Pathways

Studies employing c-fos-immunocytochemistry as a marker of cellular activation have revealed that vomeronasal neurons are activated by chemosensory stimulation. Following mating with a male, female rats exhibited a VNO-dependent increase in c-fos expression in all vomeronasal system nuclei. A particularly dense concentration of activated cells was observed in the mitral and granule layers of the AOB (Figure 3a) and in the mAMYG (Figure 3b), with increases in cellular activation also noted in the BNST, the nAOT, and the ventrolateral VMH. Female cervical stimulation or exposure of male rats to female-soiled bedding evoked a similar pattern of c-fos activation. In the male hamster, mating or exposure to vaginal discharge also produced activation of the vomeronasal system.



(a)



(b)

Figure 3 Samples from female rats exposed to 3 h of repetitive mating. a, the accessory olfactory bulb; b, the medial amygdala (mAMYG). c-Fos-immunopositive cells are indicated by small black dots. Note the cluster of c-fos-immunopositive cells parallel to the optic tract in the mAMYG. From Dudley, C. A., Sudderth, S. B. and Moss, R. L. (1992), LHRH neurons in the medial septal-diagonal band-preoptic area do not project directly to the hippocampus: a double-labeling immunohistochemical study, *Synapse* **12**, 139–146, with permission.

Activation of Gonadotropin-Releasing Hormone Neurons

During copulation, the female rat becomes more receptive over time and exhibits a release of LH (Figure 4). Because LH release is stimulated by GnRH, the chemosensory stimulation obtained during copulation is likely to activate GnRH neurons. Indeed, double-labeling immunocytochemical studies demonstrated that mating activated approximately 80% of the GnRH neurons, whereas after VNO removal only approximately 38% of the GnRH neurons were activated (see Figure 5). Similar studies

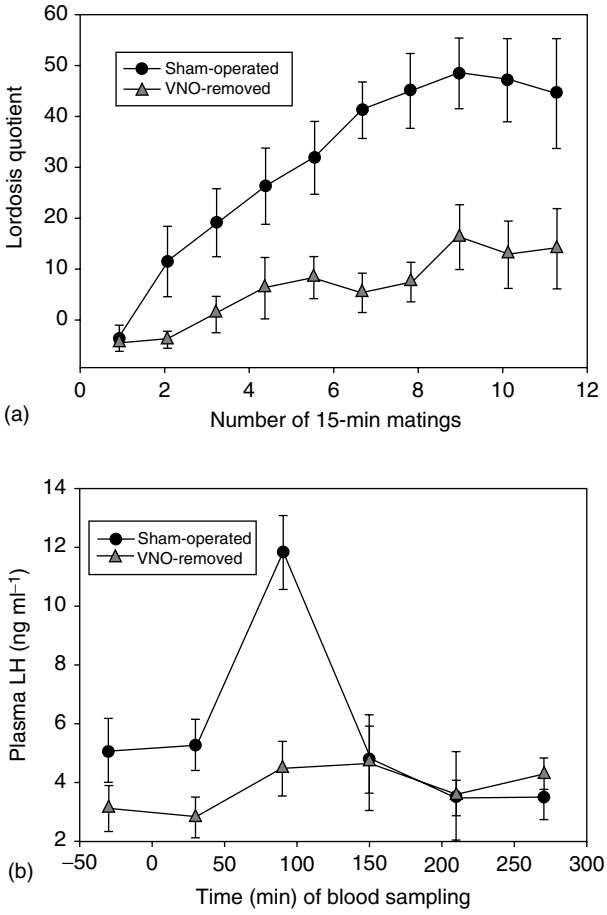


Figure 4 The effect of VNO removal in female rats. a, On lordotic responsiveness; b, on mating-induced luteinizing-hormone (LH) release. From Rajendren, G., Dudley, C. A. and Moss, R. L. (1990), Role of the vomeronasal organ in the male-induced enhancement of sexual receptivity in female rats, *Neuroendocrinology* **52**, 368–372, with permission.

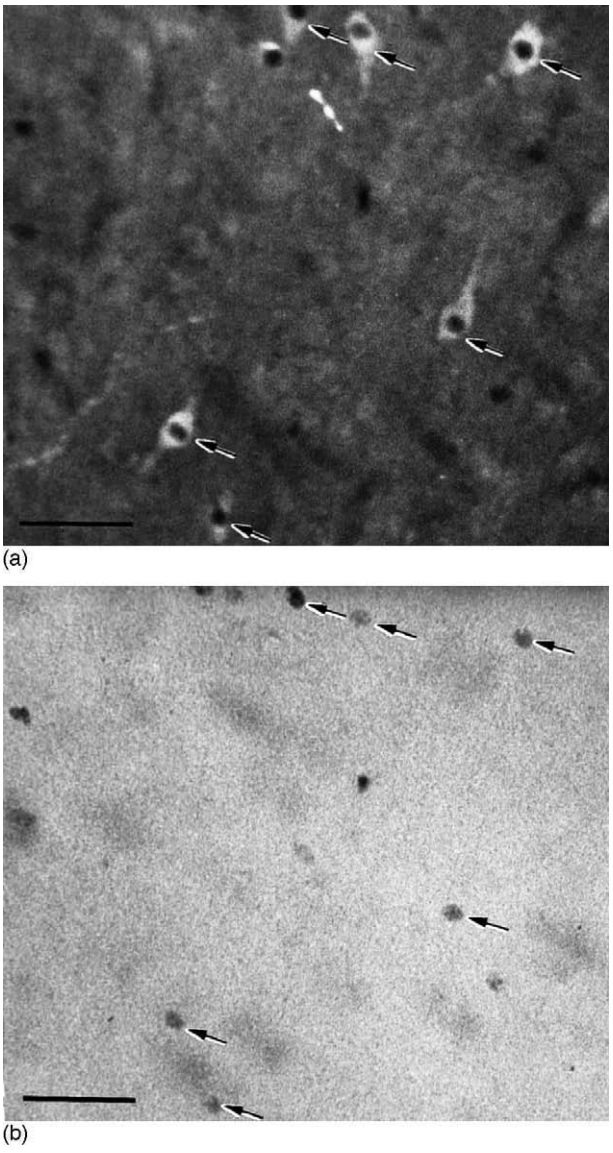


Figure 5 GnRH-containing neurons that were activated by 3 h of repetitive mating in the female rat. a, Arrows point to GnRH neurons labeled immunofluorescently with Texas Red; b, arrows point to c-fos-immunopositive nuclei. In a, the five GnRH-containing cells also contained the c-fos protein. From Rajendren, G., Dudley, C. A. and Moss, R. L. (1993), Influence of male rats on the luteinizing hormone-releasing hormone neuronal system in female rats: role of the vomeronasal organ, *Neuroendocrinology* **57**, 898–906, with permission.

from other laboratories also revealed an increase in the activation of GnRH neurons in the female mouse during mating and in the female rat during vaginocervical stimulation. These observations provide firm evidence that chemosensory stimulation at the level of the VNO influences endocrine events at the hypothalamic level.

Vomeronasal Organ Response to Olfactory Stimuli and Olfactory System Response to Pheromonal Stimuli

Recent reports have indicated that cross-talk seems to exist between the vomeronasal system (VNS) and the main olfactory system. Although the VNO is mainly dedicated to sensing pheromones and the main olfactory epithelium that lines the nasal cavity senses common odors, recent findings make us rethink how pheromones control hormonal responses. This is particularly relevant for humans, in which adults lack a functional VNO. In mammals, mating and the reproductive behavior of females are controlled by GnRH neurons in the hypothalamus. GnRH is secreted into the hypophysial portal system, through which it regulates the action of gonadotropin-secreting pituitary neurons. Recent findings have brought the main olfactory system into the limelight in the pheromone and behavior field, not only by showing its crucial role in sexual and aggressive behavior but also by revealing a direct anatomical projection from the

main olfactory epithelium to luteinizing hormone-releasing hormone (LHRH) neurons in the hypothalamus, which regulate the gonadal hormone release. There is also evidence that the VNO in the sexually naïve male hamster is crucial for the female pheromone-induced mating behavior in male, but the main olfactory system takes over the functional role of the VNO in the context of sexual behavior. Thus, the traditional distinction that common odors are perceived through the olfactory pathway and pheromones by the vomeronasal pathway does not hold true; rather, the main olfactory system and vomeronasal system complement one another in chemosensory communication.

In the human fetus, a VNO develops and appears to guide the migration of GnRH neurons to the hypothalamus. However, the lack of a functional VNO in human adults does not imply that pheromone signaling does not occur in humans. Most of the vomeronasal receptor genes in the human genome are pseudogenes; however, four of these seem to be functional, of which one, hV1RL1, is expressed in olfactory mucosa and is likely to have a transducing function. At present, the only reproducible pheromonal effect in humans is the effectiveness of axillary secretions from women in different stages of their menstrual cycles to alter the cycle lengths of other women living in close proximity, thus causing menstrual synchronization. The recent findings in mice that the main olfactory pathway is required for reproductive behavior and that it projects to and activates hypothalamic GnRH/LHRH neurons might explain how organisms without a functional VNO, such as humans, use olfactory inputs in order to activate subcortical hard-wired circuits controlling innate sexual and reproductive behavior. Advances in neuroimaging have allowed us to visualize brain activity patterns in awake humans, and these techniques have shown that sex pheromones activate wide regions of the human brain, including areas traditionally thought to mediate cognition. These studies suggest that, as in the rodent, the pheromone-activated endocrine system is much more extensive than previously thought. Thus, it appears that we need to learn more about how intimately neuroendocrine functions, controlled by pheromones acting through our noses, might interact with other operations within the brain to control human behavior and cognition.

Stress and Reproduction

Social Interactions

Stresses of a social nature have the potential to exert negative effects on reproduction. The most common social stress in animals of the same species is caused by high population density. Crowded conditions give rise to complex emotional states that may interfere with reproductive function. In addition, pheromonal cues emitted and detected in crowded conditions contribute to a decline in endocrine function. Although the exact mechanisms are uncertain, a wide array of noxious physical stimuli and distressed emotional states, including those of social origin, can enhance the activity of the hypothalamic-pituitary-adrenal (HPA) axis. The enhanced activity of this system is correlated with the suppression of reproductive processes at many levels in the reproductive axis. This suppression appears to be brought about by several specific actions that affect the GnRH system of neurons. First, corticotropin releasing hormone (CRH) and the proopiomelanocortin (POMC)-derived peptides adrenocorticotropic hormone (ACTH) and β-endorphin can act in the brain to inhibit the release of GnRH. Second, corticosteroids themselves can decrease the responsiveness of the pituitary to GnRH. All these actions can interfere with the stimulatory effect of LH on steroidogenesis in the gonad. However, the extent of reproductive inhibition varies with social rank, sex, maturity, season, and some other factors.

The negative effect of emotional stress on reproduction is well illustrated in studies performed on cages of mice that are densely populated. Typically, a large cage is populated by a few pairs and the population is allowed to grow until it regulates itself. Maintaining this regulated population level often involves a marked inhibition of sexual maturation among young animals, as well as a cessation of reproduction by adults. In crowded deer mice, the effects of high population density on reproductive function are profound. When young deer mice are allowed to grow in pairs in small cages with adults of the opposite sex, they become fertile at 5–8 weeks of age. On rearing in dense populations, however, by 12–13 weeks of age, as many as 95% of these deer mice still are not fertile. According to Terman and colleagues, many die at well over 1 year of age without ever having achieved fertility. To some extent, pheromones can alleviate the negative effects of environmental stress on estrous cycling. Female mice housed in grouped, crowded conditions experience a lack of estrous cycling, which can be overcome by urinary cues from a conspecific male or by putative pheromones isolated from male mouse urine.

Agonistic Interactions

Intermale aggression in the form of fighting is common among densely populated animals. Losers of these

fights as well as other social subordinates and most of the onlookers seem physiologically stressed. Male mice fight to establish dominance even when only two are brought together. Agonistic interactions have been shown to have a profound impact on the endocrine system. The defeated animal experiences an almost immediate inhibition of LH secretion, lowered testosterone levels, and a dramatic rise in the secretion of corticosterone. Moreover, the increase in corticosterone can be conditioned to occur in response to the mere presence of a dominant male. Thus, even without the occurrence of overt fighting, subordinate animals may have depressed LH secretion. Such alterations in hormone levels may have negative consequences for health and fertility because the long-term hyperactivation of the HPA axis is related to a decline in immune system activity and reduced gonadal activity may result in sterility.

In higher vertebrates, behavioral responses to unfamiliar conspecifics are dependent on the actual environmental situation as well as social experiences made during behavioral development. Sapolsky's studies of a population of wild African olive baboons found that high- and low-ranking baboons differ in levels of cortisol and testosterone during periods of social instability, when overt aggression is quite common. Dominant males tended to have the lowest levels of cortisol. During periods of relative stability in the colony, differences in cortisol and testosterone between dominant and subordinate males were no longer detectable. Differences in cortisol levels appear to be extremely important, not only in terms of reproductive success but also in terms of survival. Males with the lowest cortisol levels survived the longest. A strong relationship exists among environmental factors (the rearing conditions as well as the actual situation), aggressive behavior, endocrine responses, and health.

Recently, two synthetic pheromones, 2,3-dehydro-*exo*-brevicomin (DHB) and 2-*sec*-butyl-dihydrothiazole (SBT), the urinary constituents in male mice, have been shown to elicit aggression in mice. The level of two other urinary constituents from male mouse, α- and β-farnesene, were elevated in the excreted urine of dominant male mice even several days after establishing dominance. The farnesene, absent in bladder urine but abundant in the preputial gland, when mixed with water or bladder urine significantly reduced the amount of time subordinate males investigated the urine deposits of dominant males. Thus, these four different urinary constituents have been shown to contribute to male–male agonistic interactions in mice.

Nonsocial Interactions

Reproductive events are also affected by stress encountered in nonsocial situations. The endocrine system appears to be more susceptible to emotional stress than to physical stress. For example, immobilized female rats display an immediate loss of LH pulsatile activity, but no change in pulsing was observed after an experimenter administered limb fracture. The stressful nature of immobilization has also been observed in cattle restricted for the first time by a stanchion, haltered, and prepared surgically for sequential blood collection via jugular cannulae. In these conditions, greatly elevated plasma cortisol levels and no pulsatile release of LH were noticed. However, repeated exposure to this procedure resulted in low levels of plasma cortisol and a normal episodic pattern of LH release.

Stress Signaling Mediated by Pheromones

Pheromone detection is an important component in the animal's ability to respond appropriately to stressful cues in social and agonistic interactions. It is well documented in several mammalian species that animals experiencing stress and fear release chemical signals that induce behavioral, endocrinological, and immunological changes in the recipient animals of the same species. Aggressive behavior is displayed in many species in day-to-day social interactions. This behavior is elicited by the appropriate recognition of an adversary as a potential threat or sexual competitor, and this recognition involves the detection of social odors by the VNS and/or olfactory system. The involvement of the VNS in such recognition is clearly demonstrated by several studies. Genetic ablation of the TrpC2 channel results in the failure of the male mice to display aggression toward other males, similar to behavior shown by male mice after vomeronasalectomy, and also eliminates aggressive behavior of lactating female mice. However, in contrast, vomeronasalectomy does not reduce the maternal aggression of lactating female rats. Vomeronasalectomy also has been shown to decrease aggressive behavior in the male lesser mouse lemur, a prosimian.

In socially stressed crowded female rats, the removal of the VNO eliminates the resumption of estrous cycling induced by male rat urine. Mechanisms underlying the VNO-dependent responses to sociogenic stress are largely unknown at this time. Finding neural sites in which the stress-activated HPA axis interacts with the vomeronasal system seems a good strategy for determining where to begin research from a mechanistic point of view.

See Also the Following Articles

Aggressive Behavior; Crowding Stress; Reproduction, Effects of Social Stress on.

Further Reading

Boehm, U., Zou, Z. and Buck, L. B. (2005). Feedback loops link odor and pheromone signaling with reproduction. *Cell* **123**, 683–695.

Brennan, P. A. and Keverne, E. B. (2004). Something in the air?: new insights into mammalian pheromones. *Current Biology* **14**, R81–R89.

Bronson, F. H. (1988). Seasonal regulation of reproduction in mammals. In: Knobil, E. & Neill, J. D. (eds.) *The physiology of reproduction* (vol. 2), pp. 1831–1871. New York: Raven Press.

Bronson, F. H. (1995). Seasonal variation in human reproduction: environmental factors. *Quarterly Review of Biology* **70**, 141–164.

Calegro, A. E., Bagdy, G. and D'Agata, R. (1998). Mechanisms of stress on reproduction: evidence for complex intra-hypothalamic circuit. *Annals of the New York Academy of Sciences* **851**, 364–370.

Dudley, C. A., Rajendren, G. and Moss, R. L. (1996). Signal processing in the vomeronasal system: modulation of sexual behavior in the female rat. *Critical Review of Neurobiology* **10**, 265–290.

Dudley, C. A., Sudderth, S. B. and Moss, R. L. (1992). LHRH neurons in the medial septal-diagonal band-preoptic area do not project directly to the hippocampus: a double-labeling immunohistochemical study. *Synapse* **12**, 139–146.

Dulac, C. and Torello, T. A. (2003). Molecular detection of pheromone signals in mammals: from genes to behaviour. *Nature Reviews Neuroscience* **4**, 551–562.

Halpern, M. and Martinez-Marcos, A. (2003). Structure and function of the vomeronasal system: an update. *Progress in Neurobiology* **70**, 245–318.

Kumar, A., Dudley, C. A. and Moss, R. L. (1999). Functional dichotomy within the vomeronasal system: distinct zones of neuronal activity in the accessory olfactory bulb correlate with sex-specific behaviors. *Journal of Neuroscience* **19**(20), RC32.

Luo, M., Fee, M. S. and Katz, L. C. (2003). Encoding pheromonal signals in the accessory olfactory bulb of behaving mice. *Science* **299**, 1196–1201.

Marchlewska-Koj, A. (1997). Sociogenic stress and rodent reproduction. *Neuroscience and Biobehavioral Review* **21**, 669–703.

Meredith, M. and Fernandez-Fewell, G. (1994). Vomeronasal system, LHRH, and sex behaviour. *Psychoneuroendocrinology* **19**, 657–672.

Mombaerts, P. (2004). Genes and ligands for odorant, vomeronasal and taste receptors. *Nature Reviews Neuroscience* **5**, 263–278.

Moss, R. L., Flynn, R. E., Shen, X. M., et al. (1997). Urine-derived compound evokes membrane responses in mouse vomeronasal receptor neurons. *Journal of Neurophysiology* **77**, 2856–2862.

Przekop, F. and Tomaszewska, D. (1996). Responses in the hypothalamic monoaminergic system activity in ewes to beta-endorphin, CRF, and their antagonists. *Acta Neurobiologiae Experimentalis* **56**, 807–817.

Rajendren, G., Dudley, C. A. and Moss, R. L. (1990). Role of the vomeronasal organ in the male-induced enhancement of sexual receptivity in female rats. *Neuroendocrinology* **52**, 368–372.

Rajendren, G., Dudley, C. A. and Moss, R. L. (1993). Influence of male rats on the luteinizing hormone-releasing hormone neuronal system in female rats: role of the vomeronasal organ. *Neuroendocrinology* **57**, 898–906.

Yoon, H., Enquist, L. W. and Dulac, C. (2005). Olfactory inputs to hypothalamic neurons controlling reproduction and fertility. *Cell* **123**, 669–682.

Phobias *See:* Anxiety.

Photoperiodism *See:* Seasonal Changes in Stress Responses.

Physical Injury *See:* Combat, Acute Reactions to; Loss Trauma; Multiple Trauma.

Pituitary Regulation, Role of

F A Antoni
Centre for Integrative Physiology, University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F A Antoni, volume 3, pp 169–174, © 2000, Elsevier Inc.

The Endocrine Response to Stress
General Overview of the Control of the Pituitary Gland
Neurohypophysial Responses to Stress
Adenohypophysial Hormones in Stress
Intermediate Lobe

Glossary

Neuroendocrine motoneurons	Hypothalamic nerve cells that produce and secrete neurohormones into the hypophysial portal circulation and thus regulate the synthesis and release of anterior pituitary hormones.
Pituitary gland	The master gland of the endocrine system; as such, it mounts the primary endocrine response to stress, alongside the cells of the adrenal medulla, which are controlled by the autonomic nervous system.

The Endocrine Response to Stress

The stress response is characterized by marked changes of the levels of several hormones in blood. In the short term, the hormonal response to stress facilitates the mobilization of nutrients, e.g., breakdown of glycogen, promotes cardiovascular adaptation, e.g., increase of heart rate, the preservation of body fluid and salt, e.g., increase of salt reabsorption, as well as behavioral responses, e.g., explorative behavior. Exactly which hormone is responsible for a particular function may vary from species to species. For instance, the pituitary hormone prolactin is a hormone of salt-water balance in fish, while in mammals it is involved in the development of the breast in pregnancy and milk production during lactation. Of course, this latter process is a major strain on the osmotic balance of the mammalian organism, hence the phylogenetic traces of osmoregulation are evident, but only during breastfeeding and only in the female. Remarkably, prolactin is acutely released during stress in virtually all species, including male mammals; however, the functional significance of this response is still not understood.

If the stressor is chronic, i.e., recurs repeatedly or is never removed, the stress response may undergo habituation. However, in many cases this is not sufficient to prevent long-term detrimental changes of physical condition, such as reduced resistance to infectious agents, hypertension, and behavioral anomalies.

General Overview of the Control of the Pituitary Gland

The pituitary gland consists of three functionally and structurally distinct compartments (Figure 1).

The neurohypophysis or the posterior lobe is a large array of axon terminals and their supporting glial cells called the pituicytes. The neuronal perikarya that give rise to the terminal axons in the neurohypophysis lie in the hypothalamic paraventricular and supraoptic nuclei and form the magnocellular neurosecretory system (Figure 1). Separate populations of these cells produce the small peptide hormones vasopressin or oxytocin, which are synthesized in the cell bodies, transported along the axons, and released into the systemic circulation by the nerve endings in the neurohypophysis.

The adenohypophysis or anterior lobe contains various types of cells that are derived from the neural placode epithelium and produce polypeptide hormones for export into the systemic circulation. These hormones are fundamentally important for the control of the endocrine system. The cell types and their secreted hormones are as follows: corticotrophs, adrenocorticotropin (ACTH); mammotrophs, prolactin (PRL); somatotrophs, growth hormone (GH); somatomammotrophs, GH and PRL; thyrotrophs, thyroid-stimulating hormone (TSH); and gonadotrophs, luteinizing hormone (LH) and follicle-stimulating hormone (FSH). In addition, a further type of cell, the folliculo-stellate cells that are of mesenchymal origin, produce various cytokines and peptides that may be important for local paracrine control of the adenohypophysis.

The functional activity of the cells in the adenohypophysis is controlled by neurohormones (hypothalamic-releasing and -inhibiting factors) of hypothalamic origin. These are synthesized by neurons of the medial hypothalamus and secreted by axons that terminate in the external zone of the median eminence, which is a specialized neurovascular contact site at the base of the brain. This is called the parvicellular hypophysiotropic neurosecretory system of the hypothalamus (Figure 1).

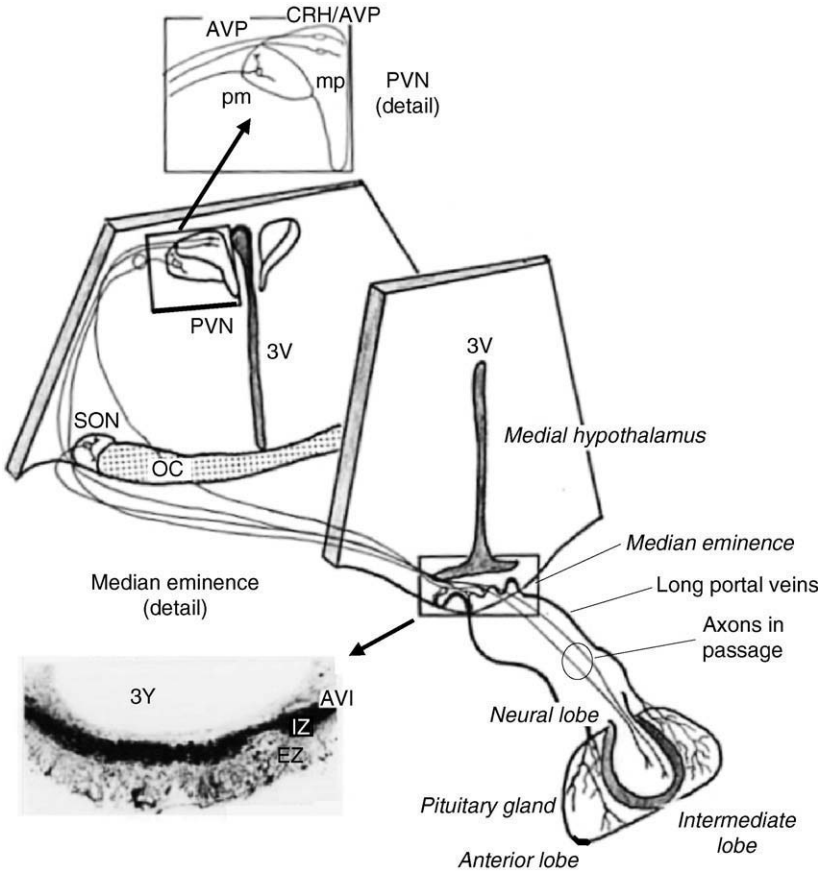


Figure 1 Schematic representation of the hypothalamo-hypophysial system that controls the secretion of ACTH. Cell bodies of parvicellular hypophysiotropic neurons synthesize corticotropin releasing hormone (CRH) and vasopressin (AVP) in the hypothalamic paraventricular nucleus (PVN). The axons reach the median eminence along an archlike pathway. The fibers first pass laterally above the fornix (F) and then turn ventral and caudal to reach the rostral part of the median eminence through the lateral retrochiasmatic area at the base of the brain, behind the optic chiasm (OC). Once in the median eminence, the fibers form terminal varicosities abutting on the pericapillary space in the external zone (EZ) of the median eminence. The effluent blood collects in the long portal veins and irrigates the anterior lobe of the pituitary gland. Magnocellular hypothalamic neurons in the PVN and the supraoptic nucleus (SON) produce AVP. The axons of these cells take a similar course in the hypothalamus as the parvicellular axons to reach the median eminence. In the median eminence, the magnocellular axons remain largely in the internal zone and appear to release AVP through conventional exocytotic mechanisms. AVP released in this manner is also found in hypophysial portal blood. The coronal section of the median eminence shown here (median eminence detail) was stained with an antiserum against vasopressin-associated neurophysin and was prepared from the brain of an adrenalectomized rat. Adrenalectomy markedly increases the AVP content in the external zone. 3V, third cerebral ventricle; pm, magnocellular part of the PVN; mp, medial parvicellular compartment of the PVN.

The capillary blood vessels in the median eminence lie immediately adjacent to the parvicellular nerve terminals and have special large pores to allow diffusion of neurohormones released by the nerve fibers into the blood irrigating the median eminence. The venous effluent of the median eminence carries the hypothalamic hormones to the adenohypophysis, where the capillaries again branch out into smaller vessels, forming the pituitary portal system that irrigates the adenohypophysis. The transit time in this system from the nerve endings to the hormone producing cells is approximately 5 s.

The internal zone of the median eminence contains the varicose nerve fibers of the magnocellular neurons running to the neurohypophysis. Functional evidence indicates that vasopressin and oxytocin are also secreted in the median eminence by these fibers in passage, thus neurohypophysial secretions may also control anterior pituitary function.

In addition to hypothalamic and paracrine pituitary-derived factors, the function of anterior pituitary cells may also be influenced by factors reaching the adenohypophysis via the systemic circulation. For instance, steroid hormones produced and secreted by the adrenal cortex influence the synthesis and the release of ACTH. Cytokines derived from the immune system may also target these cells. The actions in the adenohypophysis of factors derived

from the endocrine target glands of anterior pituitary hormones are important for the formation of neuroendocrine feedback loops.

The intermediate lobe of the pituitary gland is composed almost entirely of cells that produce melanocyte-stimulating hormone, which is important in species that change their skin color rapidly as part of their adaptive response. The intermediate lobe is innervated directly by nerve fibers originating from the medial basal hypothalamus. It has no obvious functional anatomical communication with the other two lobes.

Neurohypophysial Responses to Stress

Both vasopressin and oxytocin, the neurohormones released from the nerve terminals in neurohypophysis, have multiple functions in the body.

The principal function of neurohypophysial oxytocin is the contraction of the uterus and the ejection of milk during breastfeeding. Actions on the kidney are controversial. Oxytocin may also contribute to the regulation of prolactin secretion by the adenohypophysis, as receptors and actions on prolactin release have been described. Some stressors, e.g., immobilization, increase the plasma levels of oxytocin in experimental animals; however, the importance of this in the stress response is not understood.

Vasopressin derived from the neurohypophysis increases blood pressure by contracting blood vessels, increases water reabsorption in the kidney (hence its other name, antidiuretic hormone [ADH]), and promotes glycogenolysis in the liver. The relative importance of these effects shows considerable species variation. In humans only the actions on the kidney appear relevant.

A further important action of vasopressin in the body is the stimulation of ACTH secretion. At physiological concentrations of vasopressin, this action appears to require a cohormone called corticotropin-releasing hormone (CRH). Vasopressin derived from the neurohypophysis is distributed in the systemic circulation and reaches levels required to influence ACTH release (0.1–2 nM) only in larger mammals (sheep, cat) where peripheral levels of vasopressin may reach 1–2 nmol/L. In rodents and humans, plasma levels rarely exceed 10 pmol/L. However, vasopressin released by the fibers of magnocellular neurons in passage through the median eminence and by their dendrites in the paraventricular nucleus (Figure 1) significantly contributes to the control of ACTH secretion.

The magnocellular neurons that produce vasopressin are often referred to as the osmosensitive neurosecretory system. It is important to bear in mind that these neurons are osmosensitive in the widest sense of the word. On the one hand, changes of plasma osmolality will directly enhance the firing rate and secretory activity of these cells. On the other, conditions that potentially threaten osmotic homeostasis and body fluid volume such as strenuous exercise, pain, and hemorrhage all stimulate the release of vasopressin through neural afferent pathways.

Adenohypophysial Hormones in Stress

The two main anterior pituitary hormones released by all types of stress are ACTH and PRL. While the physiological significance of ACTH in the stress response is firmly established, the role of PRL is not well understood. The stress-induced increase of plasma PRL is, however, important in the clinic as it poses a potential problem for the differential diagnosis of hyperprolactinemia of other causes, e.g., pituitary tumors.

Hypothalamic Control of ACTH

Like most midsize (39 amino acid residues) polypeptide hormones, ACTH is initially synthesized as a larger precursor molecule called proopiomelanocortin. In the anterior pituitary gland the main hormones derived from this precursor are ACTH and β-lipotropin. In target tissues such as the adrenal cortex extracellular proteases may cleave these hormones further to other biologically active fragments such as β- and γ-melanocyte-stimulating hormone.

The main mediators of the hypothalamic control of ACTH are CRH and vasopressin. CRH is produced by parvicellular hypophysiotropic neurons located in the medial parvicellular part of the hypothalamic paraventricular nucleus (Figure 1). These parvicellular neurons also produce vasopressin, which is co-packaged with CRH and co-released from the same secretory vesicles. Significantly, the relative amount of CRH and vasopressin synthesized and released by parvicellular neurons shows wide variation. In humans the level of vasopressin in parvicellular neurons increases with age. In experimental animals, vasopressin synthesis in parvicellular hypophysiotropic neurons may be enhanced from one- to twofold that of CRH to eight- to 10-fold CRH by various forms of chronic or repeated stress as well as adrenalectomy. This is somewhat paradoxical, as chronic stressors cause persistent increases of plasma corticosteroids whereas adrenalectomy brings about a drastic reduction of these hormones. The likely explanation for these findings is that the actions of corticosteroids on hypothalamic neurons are context dependent, i.e., are a function of the state of activation of other signaling pathways that regulate

the expression of vasopressin in parvicellular hypophysiotropic neurons.

As outlined in the previous section, the magnocellular vasopressinergic system is also involved in the control of anterior pituitary function. The innervation and regulation of these cells by afferent neural pathways are distinct from that of the parvicellular CRH neurons. However, the dendrites of vasopressinergic magnocellular neurons form synaptic contacts with CRH containing parvicellular neurons, and vasopressin appears to influence the level of CRH in hypophysial portal blood. Thus, there is intrahypothalamic interaction between the magnocellular and parvicellular systems controlling the adenohypophysial secretion of ACTH. In addition, vasopressin derived from magnocellular axons in the median eminence is found in hypophysial portal blood.

A prominent functional interaction between CRH and vasopressin is evident at the level of the pituitary gland. CRH stimulates the release of ACTH through the intracellular messenger cyclic AMP. Vasopressin alone has little or no effect on ACTH release or cAMP synthesis, respectively, but synergistically enhances the effect of CRH on both ACTH release and cellular cAMP content. The basis of this is the augmentation of CRH-induced cAMP synthesis by vasopressin through protein kinase C. Thus, the adenylyl cyclase system of the adenohypophysial corticotroph cell appears to behave as a coincidence detector; it responds to vasopressin only in the presence of CRH. It is therefore likely that corticotroph cells express specialized adenylyl cyclase proteins (see **Adenylyl Cyclases and Stress Response**) that are activated by protein kinase C in the presence of the stimulatory G protein G_s. Indeed, two species of protein kinase C stimulated adenylyl cyclase (isoforms 2 and 7) have been found in the adenohypophysis, and evidence for different pathways of cAMP biosynthesis induced by CRH alone and CRH and vasopressin in combination, respectively, has been reported. There are some species variations in the potency of CRH and vasopressin to stimulate ACTH. In the sheep, vasopressin appears to be the more potent ACTH secretagogue.

The significance of vasopressin in the control of ACTH release lies in the fact that in the presence of this neurohormone corticotroph cells become relatively resistant to the negative feedback action (see below) of adrenal corticosteroids. The opposition of glucocorticoid feedback is inversely correlated to the levels of cAMP, indicating that the intracellular levels of cAMP are major determinants of the efficiency of corticosteroid feedback inhibition at the pituitary level. The concentration of vasopressin required for inducing corticosteroid resistance is in the range of 1–2 nM, which requires either vasopressin derived from magnocellular neurons or an increase of vasopressin expression and biosynthesis by parvicellular neurons such as that seen in chronic stress.

Peripheral and Paracrine Factors

The release of cytokines during inflammation appears to influence anterior pituitary ACTH release. For example, interleukin-1 induces ACTH secretion by an action in the brain and possibly directly on corticotroph cells. Of further interest is the production of leukocyte-inhibiting factor (LIF) and migration inhibitory factor (MIF) in corticotroph cells. LIF appears to be important for the maintenance of corticotroph functions, while MIF may attenuate corticosteroid feedback.

Corticosteroid Feedback Inhibition

Generally the plasma levels of adrenal corticosteroids are tightly controlled by a neuroendocrine feedback loop. Identified targets of corticosteroid action are adenohypophysial corticotrophs and parvicellular hypophysiotropic neurons that produced CRH and vasopressin. These are clearly sites of physiological relevance that control the final common elements of the neuroendocrine pathway that leads to the stimulation of ACTH. Under certain conditions, some of which are clearly pathological, glucocorticoid feedback inhibition is reduced. A well-characterized physiological example in humans is strenuous exercise. Cushing's disease and Nelson's syndrome are characterized by compromised glucocorticoid feedback at the pituitary level, which may be due to the tumoral transformation of corticotrophs. Furthermore, depressed patients have abnormally high levels of plasma corticosteroids and reduced feedback inhibition. In animals, chronic autoimmune inflammation produces high circulating levels of ACTH and corticosteroids. At the hypothalamic level this condition is characterized by a marked increase of vasopressin expression in parvicellular CRH neurons of the paraventricular nucleus. Furthermore, surgical stress and hemorrhage are also characterized by reduced corticosteroid feedback at the pituitary level, and it is known that under these conditions the magnocellular vasopressinergic neurons become active. Overall, this evidence indicates that vasopressin is mobilized in the short term as well as in the long term to modulate the efficiency of corticosteroid feedback inhibition of pituitary ACTH secretion.

Intermediate Lobe

The intermediate lobe is stimulated by epinephrine and norepinephrine derived from the adrenal medulla as well by hypothalamic CRH. It is normally inhibited

by dopaminergic input from the arcuate nucleus. The physiological role of the intermediate lobe in adult humans is unclear. In rodents, stress induces the secretion of alpha melanocyte stimulating hormone (α-MSH), which is not susceptible to corticosteroid feedback inhibition.

See Also the Following Articles

Adenylyl Cyclases and Stress Response; Vasopressin.

Further Reading

Antoni, F. A. (1986). Hypothalamic regulation of adrenocorticotropin secretion: advances since the discovery of 41-residue corticotropin-releasing factor. *Endocrine Reviews* **7**, 351–378.

Antoni, F. A. (1993). Vasopressinergic control of anterior pituitary adrenocorticotropin secretion comes of age. *Frontiers in Neuroendocrinology* **14**, 76–122.

Antoni, F. A. (1996). Calcium checks cyclic AMP – mechanism of corticosteroid feedback in adenohypophysial corticotrophs. *Journal of Neuroendocrinology* **8**, 659–672.

Antoni, F. A., Sosunov, A. A., Haunso, A., Paterson, J. M. and Simpson, J. (2003). Short-term plasticity of cyclic adenosine 3′,5′-monophosphate signaling in anterior pituitary corticotrope cells: the role of adenylyl cyclase isotypes. *Molecular Endocrinology* **17**, 692–703.

Calandra, T., Bernhagen, J., Metz, C. N., et al. (1995). MIF as a glucocorticoid-induced modulator of cytokine production. *Nature* **377**, 68–71.

Chowdrey, H. S., Larsen, P. J., Harbuz, M. S., et al. (1995). Evidence for arginine vasopressin as the primary activator of the HPA axis during adjuvant-induced arthritis. *British Journal of Pharmacology* **116**, 2417–2424.

Dallman, M. F., Akana, S. F., Scribner, K., et al. (1992). Stress, feedback and facilitation on the hypothalamo-pituitary-adrenal axis. *Journal of Neuroendocrinology* **4**, 517–526.

de Kloet, E. R., Vreugdenhil, E., Oitzl, M. S. and Joëls, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* **19**, 269–301.

Engelmann, M., Landgraf, R. and Wotjak, C. T. (2004). The hypothalamic-neurohypophysial system regulates the hypothalamic-pituitary-adrenal axis under stress: an old concept revisited. *Frontiers in Neuroendocrinology* **25**, 132–149.

Jones, M. T. and Gillham, B. (1988). Factors involved in the regulation of adrenocorticotropic hormone/β-lipotropic hormone. *Physiology Reviews* **68**, 743–818.

Petrides, J. S., Mueller, G. P., Kalogeras, K. T., et al. (1994). Exercise-induced activation of the hypothalmaic-pituitary-adrenal axis: marked differences in the sensitivity to glucocorticoid suppression. *Journal of Clinical Endocrinology and Metabolism* **79**, 377–383.

Plotsky, P. M. (1991). Pathways to the secretion of adrenocorticotropin: a view from the portal. *Journal of Neuroendocrinology* **3**, 1–9.

Tilders, F. J. H., Berkenbosch, F., Vermes, I., Linton, E. A. and Smelik, P. G. (1985). Role of epinephrine and vasopressin in the control of the pituary-adrenal response to stress. *FASEB Journal* **44**, 155–160.

PMS *See:* Premenstrual Dysphoric Disorder.

Police, Stress in

R J Burke
York University, Toronto, Ontario, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R J Burke, volume 3, pp 175–178, © 2000, Elsevier Inc.

Work Stressors
Psychological Burnout
Organizational Initiatives
Individual Initiatives

Glossary

Burnout	Response to chronic work stress characterized by exhaustion, cynicism, and low feelings of accomplishment.
Coping	Responses by individuals to manage work stressors.

Work Stressors

In comparison with other occupations, police work has been identified as a particularly stressful occupation. Negative aspects of the job such as boredom,

lack of respect from members of the public, excessive paperwork, contacts with the public that are sometimes negative and confrontational, shift work, threats of violence, and the militaristic nature of the bureaucratic structure of policing are among the work stressors that confront police officers. As a result of these and other stressful aspects of policing, a variety of symptoms and reactions may occur. These include deteriorating work performance (absenteeism, low morale), negative psychological states (emotional burnout, frustration, depression, anger), and psychosomatic and physical conditions (headaches, ulcers). The stresses of policing have been offered as one reason why capable police officers leave the profession for other careers.

Figure 1 shows the wide variety of potential stressors in police work. These include external demands and factors in the broader social context, aspects of the criminal justice system, the police organization and the police role itself, aspects of individual differences, and work–family issues. In addition, individual characteristics of police officers (e.g., authoritarian personality) may increase levels of work stress. Some of these stressors are common to a number of occupations while others are unique to police work.

The police organization and the police job pose particular demands. Police officers, particularly early in their careers, work varying shifts. In addition, they often are on-call. Overtime work is an inevitable part of their job. These factors result in heavy workloads in terms of hours per week spent working.

The police role also can be a concern. Police officers are required to be suspicious and hypervigilant. They are trained to exercise emotional control. Finally, their role conveys considerable authoritativeness. Police officers are always right and have considerable power and authority over citizens.

The peer culture and socialization influences in policing are also somewhat unique. These include the presence of masculine attitudes and values (e.g., control, dominance, authority), an extreme in-group/out-group mentality, and intense solidarity with peers. Police are likely to socialize with other police officers. Since many contacts with citizens are negative in nature, this fosters a mentality of us versus them.

Cynicism among police officers has received considerable attention as well. Officers begin their careers with positive attitudes toward their jobs and their role. These attitudes have been shown to steadily deteriorate as officers approach mid-career, then

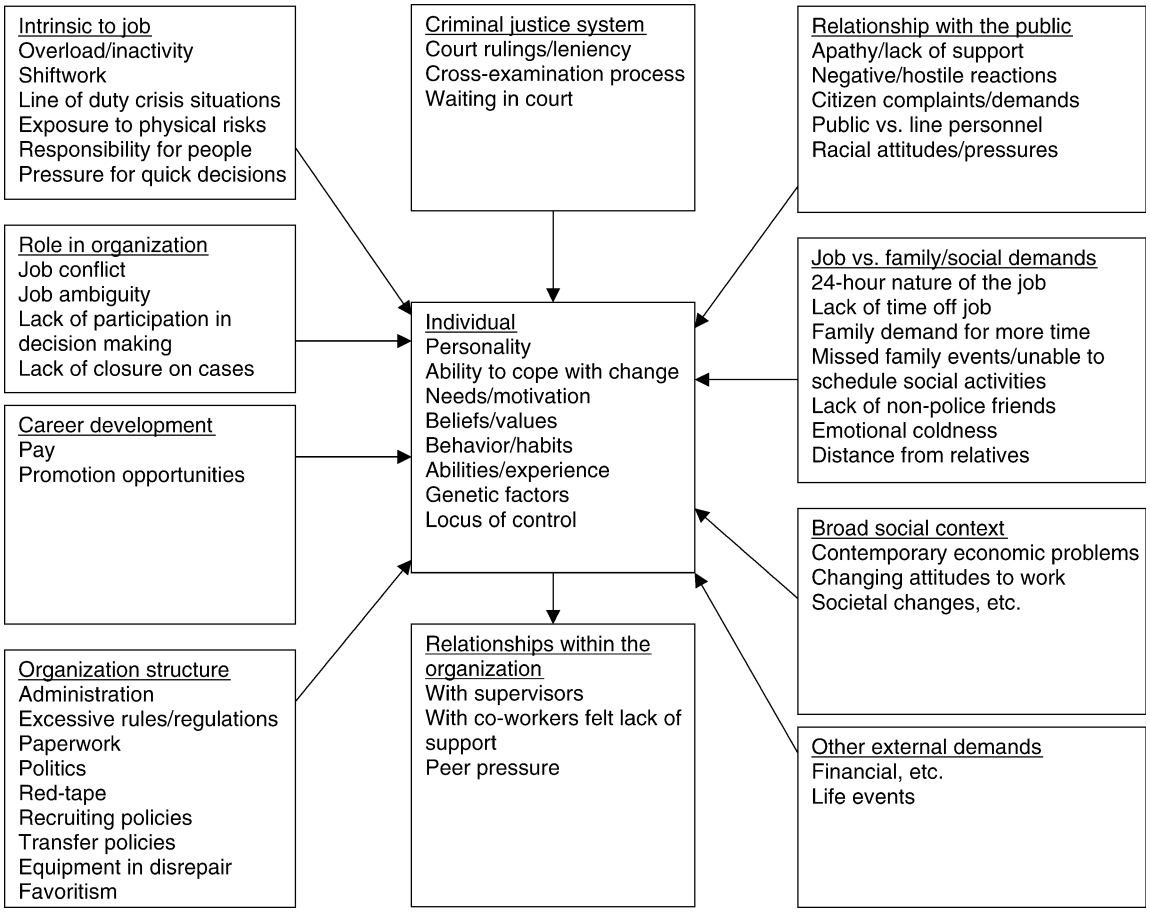


Figure 1 Understanding stress in policing.

improve modestly with increasing police tenure as officers approach retirement. These attitude changes are important to both understand and manage since police officers are vulnerable to career plateauing given the steep organizational hierarchy in police management. The challenge for senior police management is to maintain positive attitudes among officers who will never be promoted.

These factors produce some common problem areas. Police officers become overinvolved with their force and underinvolved with their spouses or partners. The cynicism police have about their profession and the broader justice system spills over to the home front. Police officers make authoritative demands at home, as they do in their work. The irregular schedule and the leave time demands results in limited time with their spouses and families. Spouses or partners often have fear or anxiety about the safety of the officer. The spouses or partners are often jealous of the time that police officers spend with coworkers. Police families are sometimes isolated from the broader community. Finally, there are weapons in the home.

Work stressors in policing have been found to be related to work–family conflict, which in turn has been related to psychological burnout. Jackson and Maslach collected data from men and women in police work and their spouses or partners. They found that police officers' levels of burnout affected their interactions with their spouses and children. Burke examined antecedents and consequences of work–family conflict among 828 men and women in police work. Ratings of work–family conflict were weakly correlated with demographic antecedents but strongly correlated with work setting characteristics, levels of social support, levels of work and non-work stress, and several outcome measures (including burnout).

The combination of the types of individuals attracted to police work and the nature of the job has been suggested to increase particular problematic behaviors. These behaviors include anger, verbal outbursts, physical violence, alcoholism, and infidelity. Forty percent of police officers reported at least one episode of physical aggression during a conflict with their spouse or partner in the previous year.

Some published literature has downplayed the negative effects on officers of participating in risky calls and has pointed to the organizational structure as the greatest source of stress and pressure, adversely affecting job satisfaction and the perceived quality of life.

Although most police officers are men, an increasing number of women have become officers over the past decade. About 15 to 20% of officers in the developed world are women. Women and men in police work generally report similar levels of work stressors, job satisfactions, burnout, and psychological well-being. Interestingly, the biggest source of job stress reported by women officers comes primarily from their male colleagues and to a lesser extent from male members of the general public. Women officers report prejudice, discrimination, and sexual harassment from both of these sources.

Psychological Burnout

Research attention has increasingly focused on psychological burnout, a common reaction to stress among police officers. Maslach and Leiter defined burnout as a syndrome of emotional exhaustion, depersonalization, and low personal accomplishment. Jackson and Maslach studied the effects of levels of self-reported psychological burnout on police officers' interactions with their spouses and children. Police officers reporting higher levels of psychological burnout were more likely to display anger, spend time off away from the family, be uninvolved in family matters, and have unsatisfactory marriages.

Empirical studies of psychological burnout among police officers have attempted to identify both individual difference characteristics and job conditions associated with higher burnout levels. Relationships of burnout with individual demographic characteristics are sometimes found but tend to be both inconsistent and weak. However, there is considerable evidence that work setting characteristics, particularly chronic work stressors, influence levels of psychological burnout. Such characteristics have included features of the job itself, quality of supervision, unmet expectations, and constraints in one's organizational environment.

Police forces are concerned when their officers use excessive force toward citizens. Recent research has shown a relationship between officer burnout levels and both attitudes toward the use of force and the actual use of force. Burned-out officers were more likely to advocate and use force and violence. Cynicism toward the general public lowers the threshold for using force or violence. Exhaustion and low professional efficacy reduce problem-solving skills while increasing use of force and violence.

Organizational Initiatives

Recruitment and Selection Interventions

Police work is a relatively high-paying occupation. As a result, a large number of individuals are interested in developing careers in policing. It seems desirable for police forces to use realistic job previews so that applicants have a greater understanding of the nature

and demands of the occupation, make greater use of personality assessment to weed out applicants likely to have a difficult time performing effectively, and use the training period to both socialize desired behaviors and identify applicants that may not perform well.

Counseling Interventions

It is critical to offer trauma counseling to police officers in cases of acute stress (using a weapon, shooting a suspect, death of a citizen in an accident). In addition, the provision of counseling to police officers via Employee Assistance Programs (EAPs) to handle alcohol abuse, work–family difficulties, and family problems is useful. The macho nature of the police culture and the real (or imagined) stigma of seeking help must be addressed.

Training Interventions

Two areas of training are evident. One specifically addresses work stressors and coping, particularly the use of active coping responses. Although such training has not been found to produce lasting benefits, it may still be justified. A second aspect of training, aimed at the organizational level, involves management development. It is unrealistic to expect that the nature and culture of policing will change dramatically or quickly, but efforts in this direction must be made.

Work–Family Initiatives

As noted earlier, the work–family interface can be particularly problematic in policing. Several work–family initiatives have been undertaken by innovative focus. These include a police spouse/partner support group, a couple workshop that both educates and addresses concrete issues, the inclusion of work–family content in training and development activities, and the use of peer counseling efforts.

Individual Initiatives

Although coping is a critical individual response to perceived work stressors due to its complexity, it is still not well understood. A simple taxonomy is to classify individual coping responses as active (problem solving, talk with others, anticipate upcoming demands) or passive (do nothing, avoidance, use alcohol). Hurrell and Kroes reported that 25% of all police officers are dependent on alcohol as a stress reliever. In a large sample of police officers, Burke found that use of active coping responses was associated with greater job satisfaction, and use of passive coping responses was associated with greater psychological distress. Use of active and passive coping responses was, not surprisingly, negatively correlated. Police officers making greater use of active coping made lesser use of passive coping responses.

See Also the Following Articles

Work–Family Balance; Workplace Stress.

Further Reading

Abdollahi, K. M. (2002). Understanding police stress research. *Journal of Forensic Psychology Practice* **2**, 1–24.

Brown, J. and Heidenshohn, F. (2000). *Gender and policing: comparative perspectives*. Basingstoke: MacMillan.

Burke, R. J. (1997). Work and non-work stressors and well-being among police officers: the role of coping. *Anxiety, Stress & Coping* **11**, 345–362.

Cooper, W. H. (1982). Police officers over career stages. *Canadian Police College Journal* **6**, 93–112.

Hart, P. M., Wearing, A. J. and Headey, B. (1993). Assessing police work experiences: Development of the police daily hassles and uplift scales. *Journal of Criminal Justice* **21**, 553–572.

Jackson, S. E. and Maslach, C. (1982). After-effects of job-related stress: families as victims. *Journal of Occupational Behavior* **3**, 63–77.

Kop, N. and Euwema, M. C. (2001). Occupational stress and the use of force by Dutch police officers. *Criminal Justice and Behavior* **28**, 631–652.

Kroes, W. H. (1976). *Society's victim – the policeman. An analysis of job stress in policing*. Springfield, IL.: Charles C. Thomas.

Kroes, W. H. and Hurrell, J. J. (1975). *Job stress and the police officer*. Washington, D.C.: U.S. Government Printing Office.

Maslach, C. and Leiter, M. P. (1997). *The truth about burnout*. San Francisco: Jossey-Bass.

Neidig, P., Russell, H. and Seng, A. (1992). Interpersonal aggression in law enforcement families: a preliminary investigation. *Police Studies* Spring 30–38.

Posttraumatic Stress Disorder – Clinical

N C Feeny and L R Stines
Case Western Reserve University, Cleveland, OH, USA
E B Foa
University of Pennsylvania, Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Overview of Posttraumatic Stress Disorder
Cognitive-Behavioral Treatments for Posttraumatic Stress Disorder
Real-World Treatment Considerations
Childhood Sexual Abuse-Related Posttraumatic Stress Disorder
Conclusion

Glossary

Arousal	A cluster of PTSD symptoms that includes difficulty concentrating, sleep disturbance, irritability and anger, exaggerated startle response, and overalertness.
Avoidance	A cluster of PTSD symptoms that includes behavioral and cognitive avoidance (i.e., avoidance of thoughts, feelings, and reminders of the trauma), emotional numbing, loss of interest in activities, feeling disconnected from others, psychogenic amnesia, and a sense of foreshortened future.
Cognitive-behavioral treatments	A class of psychosocial treatment approaches focused on the interrelationships among thoughts, behaviors, and emotions.
Posttraumatic stress disorder (PTSD)	An anxiety disorder that develops in response to a traumatic event and is characterized by three clusters of symptoms: reexperiencing, avoidance, and physiological hyperarousal. PTSD is diagnosed after a 1-month period of sustained problems following the traumatic event.
Reexperiencing	A cluster of PTSD symptoms that includes reexperiencing the traumatic event through intrusive distressing thoughts, nightmares, flashbacks, and intense emotional upset and physiological arousal on exposure to reminders of the trauma.
Trauma	An event that a person experiences, witnesses, or learns about that involves actual or threatened physical injury or death and to which the person responds with feelings of fear, helplessness, or horror.

Overview of Posttraumatic Stress Disorder

Prevalence of Traumatic Events

According to the *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.; DSM-IV) a trauma is an event that involves actual or perceived physical threat or injury, to which a person responds with fear, helplessness, or horror. Although in the past, trauma was conceptualized as unusual and outside the realm of human experience, large epidemiological studies have demonstrated high rates of trauma exposure among adults. In a nationally representative sample of 6000 U.S. adults, Kessler and colleagues found that 60% of men and 51% of women reported experiencing at least one traumatic event in their lifetime. Similar findings emerged from Norris's 1992 study using a large, racially diverse sample – 74% of men and 65% of women reported experiencing at least one traumatic event. A lower prevalence of trauma was found in a sample of members of a health maintenance organization – 43% of men and 39% of women experienced a trauma in their lifetime. Taken together, these studies suggest that trauma is very common among adults in the United States.

Diagnostic Criteria of Posttraumatic Stress Disorder

The constellation of psychological difficulties that is observed most often following a trauma is called posttraumatic stress disorder (PTSD). PTSD is an anxiety disorder, and its symptoms fall into three clusters: reexperiencing the traumatic event (e.g., intrusive and distressing thoughts, flashbacks, and nightmares), avoidance of trauma-related reminders (e.g., avoiding thoughts or situations related to the trauma, emotional numbing, sense of foreshortened future), and hyperarousal (e.g., sleep disturbance, difficulty concentrating, irritability, and hypervigilance). In order to meet diagnostic criteria for PTSD, these difficulties must last for more than 1 month and cause substantial impairment in social or occupational functioning.

Prevalence and Course of Posttraumatic Stress Disorder

In the general population, the lifetime prevalence of PTSD is estimated to be around 9%. Estimates suggest that women are two times more likely to develop PTSD than men (e.g., 10% of women vs. 5% of men

in the general population). Because a diagnosis of PTSD requires an exposure to a traumatic event, rates of PTSD are inherently higher among those who have experienced such an event, with lifetime prevalence estimated at 25%. Among trauma survivors, the prevalence of current PTSD varies quite widely. For example, 15% of Vietnam combat veterans, 12–65% of female assault survivors, and up to 40% of survivors of serious motor vehicle accidents have been reported to meet diagnostic criteria for PTSD. Thus, although trauma exposure is quite common, persistent psychological difficulties in the aftermath of a traumatic event are less so and vary in frequency, in part, as a function of the traumatic experience.

In the initial days and weeks following a trauma, it is common to experience difficulties such as fear, disrupted sleep, and difficulty concentrating. Although most people experience such difficulties shortly after a traumatic event, the majority also experience a natural reduction in these difficulties over the following several months. Some trauma survivors, however, continue to experience significant PTSD symptoms for months and even years after exposure to the traumatic event. After 1 year, chronic PTSD is unlikely to remit without treatment; thus there is a critical need for effective treatments.

Cognitive-Behavioral Treatments for Posttraumatic Stress Disorder

Several cognitive-behavioral therapies for PTSD have been developed and tested in prospective randomized studies. These treatments include prolonged exposure therapy (PE), stress inoculation training (SIT), cognitive therapy (CT), cognitive processing therapy (CPT), and eye movement desensitization and reprocessing (EMDR). In this section, these treatments are reviewed and selected outcome studies that evaluate the efficacy of these treatments are presented.

Prolonged Exposure Therapy

Exposure therapy has been extensively evaluated as a treatment for persistent, pathological anxiety. The chief aim of this form of treatment is to help patients confront feared and avoided objects, situations, memories, and images until anxiety declines. PE for trauma survivors with PTSD is a treatment program that is based on the notion that avoidance of trauma-related memories and external reminders interferes with recovery from the traumatic event. Specifically, avoidance prevents opportunities to emotionally process the trauma memory, to learn to distinguish safe from unsafe situations, and to disconfirm erroneous trauma-related cognitions, thus maintaining the unrealistic cognitions underlying PTSD that the world is utterly dangerous and that the trauma survivor is extremely incompetent in coping with stress. This view of the mechanisms involved in exposure therapy is central to emotional processing theory.

PE includes two forms of exposure: imaginal exposure and *in vivo* exposure. Imaginal exposure involves repeatedly recounting the traumatic memory. Specifically, the patient is instructed to vividly imagine the traumatic event and to describe it in detail aloud, along with the thoughts and feelings that occurred during the event. Imaginal exposure is typically conducted during several treatment sessions and via homework of listening to audiotapes of the in-session recounting of the memory. *In vivo*, or real-life, exposure involves systematic and repeated confrontation with safe or low-risk trauma-related reminders that evoke excessive or unrealistic anxiety. *In vivo* exposure is typically conducted outside of the treatment sessions as homework, during which the patient confronts situations, places, objects, or activities that are realistically safe but that trigger trauma-related fear and anxiety because they are related to the trauma. With *in vivo* exposure, the patient is encouraged to remain in the anxiety-provoking situation until his or her fear decreases by a significant amount. Other components of PE include psychoeducation about common reactions to traumatic events and breathing retraining.

Across several studies in different centers, PE has consistently been found to be highly effective. PE has demonstrated outcomes superior to supportive counseling (SC) and a wait-list control (WL). Most patients who receive PE experience a significant reduction in PTSD, depression, anxiety, anger, and guilt and experience improvements in overall functioning. In fact, in the treatment guidelines developed by experts in the field in conjunction with International Society for Traumatic Stress Studies (ISTSS), PE was identified as the most empirically supported intervention for PTSD.

Stress Inoculation Training

SIT was among the first of the cognitive-behavioral treatments to be applied to PTSD. According to the theory underlying SIT, stress occurs when people experience environmental events as exceeding their coping resources and thereby as threatening to their safety or welfare. Anxiety is a normal response to stress and signals people to increase efforts to cope with or somehow manage the stress-eliciting situation. In the case of individuals with PTSD, the anxiety evoked by trauma-related reminders is excessive and disruptive. Thus, SIT involves the learning and

repeated practice of specific coping skills, with the goal of helping patients to develop or enhance their ability to manage stress effectively and to reduce anxiety. Coping skills typically include tools such as breathing and relaxation training, guided self-dialog, assertiveness training, role-playing, covert modeling, and cognitive restructuring.

The efficacy of PE, SIT, SC, and WL were compared for female sexual-assault survivors with chronic PTSD. PE and SIT showed significant pre- to posttreatment reductions in reexperiencing and avoidance symptoms, whereas SC and WL did not. Also, at the end of treatment, 50% of patients who had SIT and 40% who had PE no longer met the criteria for PTSD; in contrast, only 10% of patients who had SC and none who had WL lost their diagnosis. At follow-up, there was a tendency for patients in the PE group to show further improvement in PTSD symptoms, whereas patients in the SIT and SC groups did not.

In a second study, the efficacy of PE and SIT was compared to the efficacy of a combination of both treatments (PE/SIT) and WL. All three active treatments resulted in considerable symptom reduction and were more effective than WL. However, PE appeared superior to the other treatments, leading to greater reductions of anxiety and depression, with fewer dropouts than both SIT and PE/SIT. At the end of treatment, 70% of patients who had PE, 58% who had SIT, and 54% who had PE/SIT lost their PTSD diagnosis compared to none in the WL group. At follow-up, women who received PE exhibited better social functioning than did those who received SIT or PE/SIT.

Cognitive Therapy

The aims of CT or cognitive restructuring (CR) for PTSD patients are to help patients to understand the role of their beliefs and appraisals of situations in influencing their emotional reactions, to identify trauma-related irrational thoughts or beliefs that trigger avoidance and/or excessive negative emotions (e.g., fear, shame, and rage), and to learn to challenge unrealistic beliefs and expectations in a rational evidence-based manner. In challenging these beliefs, evidence is weighed and alternative ways of viewing the situation are evaluated. In treatment sessions, and in daily life, patients practice responding to automatic thoughts and interpretations by reviewing the facts, considering alternative explanations, and sometimes experimenting with different ways of behaving in response to situations or events that elicit anxiety and other negative emotions. As a consequence, patients learn to evaluate whether their trauma-related beliefs and expectations accurately reflect reality and are appropriate or helpful and modify them accordingly.

In a 2005 randomized trial, PE was compared with a program that included PE and CR (PE/CR) in female victims of sexual and nonsexual assault. The results suggest that both PE and PE/CR produce substantial improvements on multiple outcomes, including evaluator ratings of PTSD severity, depression, and social functioning. Similar to the PE/SIT findings, combining PE with CR did not improve the efficacy of PE. In an examination of PE in a sample of mixed trauma victims with chronic PTSD, Marks et al. compared PE to CR, PE/CR, and a relaxation control condition (R). The results suggest that PE, CR, and PE/CR are all quite effective and superior to R. Again, both PE and CR were effective, and combining them did not improve efficacy.

Cognitive Processing Therapy

Another cognitive behavioral therapy program for PTSD is CPT. CPT was originally developed for use with sexual assault victims and comprises elements of both CT and PE. Specifically, CPT includes CR around beliefs in five core areas – safety, power and control, esteem, trust, and intimacy – written exposure (i.e., writing about the traumatic memories and reading the account aloud to the therapist).

Resick et al. compared the efficacy of 12 sessions of CPT, 9 sessions of PE, and WL in rape victims with chronic PTSD. Results indicated that after 9 weeks of PE and 12 weeks of CPT, both treatments were highly effective in reducing PTSD; posttreatment only 19.5% of completers of CPT and 17.5% of PE met the criteria for PTSD. Among the completers, 76% of patients who had CPT and 58% of patients who had PE met the criteria for good end-state functioning (i.e., low PTSD and depression). Gains were maintained over time as well. At a 3-month follow-up, only 16.2% of those who received CPT and 29.7% of those who received PE were PTSD positive. Finally, at a 9-month follow-up, 80.8% of CPT and 84.6% of PE patients remained in sustained remission from PTSD symptoms. In sum, CPT appears to be an efficacious intervention for the short- and long-term treatment of PTSD. This study also provides further evidence for the efficacy of PE.

Eye Movement Desensitization and Reprocessing

EMDR emerged as a treatment for PTSD in the early 1990s, and since that time, a number of studies have investigated its efficacy. A core component of this intervention is theorized to be the therapist's repeated elicitation of rapid, saccadic eye movements (or other bilateral stimulation) from the patient during

the processing of traumatic memories. Shapiro theorized that rapid eye movements in some way override or reverse neural blockage or obstruction induced by the traumatic event. During EMDR treatment sessions, patients are asked to generate an image of the trauma and to focus on trauma-related thoughts, feelings, and/or sensations. Simultaneously, the therapist elicits the saccadic eye movements by having patients visually track a finger rapidly waved back and forth in front of their face. Other forms of laterally alternating stimuli (e.g., tapping or alternating sounds) are sometimes used rather than the original finger tracking. Patients are asked to evaluate the cognitions associated with these images and experiences and to generate alternative cognitive appraisals of the trauma or their behavior during the event. As patients focus on the distressing images and thoughts, and later focus on the alternative cognition, the saccadic eye movement (or another form of bilateral stimulation) is intermittently generated.

EMDR has been the focus of considerable controversy due to early claims regarding remarkable success in only a single session. Of the studies evaluating the efficacy of EMDR, more recent studies have typically used well-controlled designs and thus yielded clearly interpretable results. One well-controlled 1997 study evaluated the efficacy of EMDR relative to WL for PTSD in female rape victims. Three sessions of EMDR resulted in greater improvement of PTSD symptoms (57% reduction in PTSD at posttreatment) than WL (10% reduction at posttreatment). However, in contrast to other controlled studies, one therapist conducted all the treatments, and therefore the contribution of therapist and treatment effects are confounded. A direct comparison between EMDR and a combined treatment of PE plus stress inoculation training (called trauma treatment protocol, TTP) was conducted by DeVilly and Spence. Both TTP and EMDR reduced PTSD severity (63 and 46%, respectively), but TTP patients maintained their gains at follow-up, whereas EMDR patients showed higher rates of relapse (symptom reduction at follow-up of 61 and 12%, respectively). Rothbaum, Astin, and Marsteller compared PE, EMDR, and WL in female sexual assault victims. Compared to WL, both treatments resulted in a significant reduction in PTSD severity and related psychopathology. Immediately after treatment, only 5% of PE participants and 25% of EMDR participants continued to meet PTSD diagnostic criteria, compared to 90% of WL participants. The active treatments were significantly different from WL but did not differ from one another. At follow-up, however, significantly more participants receiving PE (78%) achieved good end-state functioning than did participants receiving EMDR (35%). Taylor and colleagues compared PE, EMDR, and R and found that PE produced significantly greater reductions in reexperiencing and avoidance symptoms than EMDR and R, which did not differ from one another. Indeed, a 2001 meta-analysis found EMDR to be no more effective than exposure therapy programs and, moreover, suggested that the eye movements integral to the treatment and its underlying theory are unnecessary.

In summary, several cognitive-behavioral therapy programs have received empirical support for their efficacy in reducing PTSD severity and diagnosis in controlled studies: PE, SIT, CT, CPT, and EMDR. Studies that have attempted to combine these empirically supported approaches into new treatment packages to improve outcome have failed to demonstrate superior outcomes compared to the single-treatment protocols in these clinical trials.

Real-World Treatment Considerations

Treatment Tolerability and Preference

It also important to explore the tolerability and acceptability of empirically supported treatments for PTSD. Notably, although concerns have been raised by some clinicians and researchers that PE is too difficult to tolerate and may lead to symptom worsening and treatment drop-out, there is evidence that PE is tolerable. For example, Foa and colleagues examined self-reported symptom exacerbation in a sample of women with chronic assault-related PTSD who were receiving PE. After the first imaginal exposure, only a minority of participants in the study exhibited reliable symptom exacerbation during treatment: 10.5% reported an increase in PTSD symptoms, 21.1% reported an increase in anxiety, and 9.2% reported an increase in depression. Significantly, patients who reported such symptom exacerbation benefited from PE as much as those who did not report such exacerbation and were not more likely to drop out of treatment. Thus, although some patients experienced a brief period of exacerbation of symptoms during PE, this exacerbation decreased within 2 weeks and was unrelated to treatment outcome or completion. With respect to the completion of treatment, additional evidence supports that PE is tolerable. A 2003 review of dropout rates from various cognitive-behavioral therapy treatments for PTSD across 25 controlled studies indicated no significant differences in dropout rates among trials of PE, CT, SIT, EMDR, and combined treatments. PE did not produce a higher rate of treatment dropouts than other PTSD treatments; thus, concerns about excessive treatment dropout from PE appear to be unwarranted.

In fact, recent data demonstrate that treatments including PE are actually preferred over other psychosocial and psychopharmacological treatments. In a forced-choice comparison of treatment preference between PE and sertraline (an antidepressant medication) for PTSD, approximately 87% of female college students chose PE compared to only 7% who chose medication (6% chose no treatment). Consistent with these data, Tarrier and colleagues found that adults strongly prefer PE for PTSD compared to other psychosocial interventions, including EMDR, group therapy, and psychodynamic psychotherapy.

Treatment Dissemination

Disseminating empirically supported treatments such as PE to community-based clinicians is an important step in bridging the critical gap between research and practice for trauma survivors with PTSD. In one of the first studies to evaluate such dissemination, Foa and colleagues trained clinicians at a community-based rape crisis center in the delivery of PE for assault survivors with PTSD. These community clinicians provided treatment to a subset of participants in a randomized controlled trial, and their outcomes were compared to participants treated by those with expertise in PE at an academic medical center. After an initial training with follow-up consultation and supervision, results showed that the community mental health counselors obtained outcomes comparable to the university-based providers. This provides good evidence that empirically supported treatments such as PE can be successfully disseminated.

Childhood Sexual Abuse-Related Posttraumatic Stress Disorder

Despite the documented efficacy of PE for PTSD, there is some concern reported in the clinical literature that adults who have PTSD secondary to maltreatment during childhood, specifically child sexual abuse, may be a particularly vulnerable population for whom PE may be too difficult to tolerate. In response to this concern, a new treatment protocol, skills training affect and interpersonal regulation plus modified prolonged exposure (STAIR/MPE), was designed as a two-phase treatment, with phase 1 consisting of skills training in affect and interpersonal regulation and phase 2 consisting of modified PE (imaginal exposure plus coping skills training, with no *in vivo* exposure). Data suggested that STAIR/MPE resulted in significant reductions in PTSD and depression compared to WL. However, a critical weakness of this study is that this combination treatment was not compared to the usual PE. It is of note that two recent clinical trials demonstrated that cognitive-behavioral treatments are effective for reducing PTSD among child sexual abuse survivors without a skills-building component. Moreover, a subanalysis of the data from Foa et al.'s 2005 study demonstrated that patients whose PTSD was related to child sexual abuse benefited from PE and PE/CR as much as did those whose PTSD was related to adult trauma. Thus, it seems that the skills-building component, which extends the number of treatment sessions and may delay recovery time, is unnecessary above the brief treatments that have empirical support.

Conclusion

Although exposure to potentially traumatic events is quite common, only a small minority of people develop persistent psychological difficulties such as PTSD in response to a trauma. The most validated types of treatment for PTSD are cognitive-behavioral therapies, and among the cognitive-behavioral therapy approaches available, PE has the strongest empirical support. Further, recent data suggest that PE is a preferred treatment among those who are given various treatment descriptions for PTSD. Concerns over the tolerability of PE, specifically that PE may lead to symptom exacerbation or premature dropout, are not supported by data. STAIR/MPE, a modified version of PE, has been shown to be helpful for childhood sexual abuse survivors with PTSD. However, this modified treatment has not been compared directly to already validated treatments for PTSD such as PE. Thus, at this point, we cannot determine whether it offers incremental benefit above such treatments. Future research should continue to evaluate ways to enhance treatment outcome and to disseminate efficacious treatments to a larger number of those who suffer with PTSD.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Child Sexual Abuse; Cognitive Behavioral Therapy; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Neurobiological basis for.

Further Reading

American Psychiatric, Association (1994). *Diagnostic and statistical manual of mental health disorders* (4th edn.). Washington, DC: American Psychiatric Association.

Cloitre, M., Koenen, K. C., Cohen, L. R., et al. (2002). Skills training in affective and interpersonal regulation

followed by exposure: a phase-based treatment for PTSD related to childhood abuse. *Journal of Consulting and Clinical Psychology* **70**(*5*), 1067–1074.

Foa, E. B., Dancu, C. V., Hembree, E. A., et al. (1999). A comparison of exposure therapy, stress inoculation training, and their combination for reducing posttraumatic stress disorder in female assault victims. *Journal of Consulting and Clinical Psychology* **67**(2), 194–200.

Foa, E. B. and Rothbaum, B. O. (1998). *Treating the trauma of rape*. New York: Guildford Press.

Kessler, R. C., Sonnega, A., Bromet, E., et al. (1995). Posttraumatic stress disorder in the national comorbidity survey. *Archives of General Psychiatry* **52**, 1048–1060.

Zoellner, L. A., Feeny, N. C., Cochran, B., et al. (2003). Treatment choice for PTSD. *Behaviour Research & Therapy* **41**(8), 879–886.

Posttraumatic Stress Disorder – Neurobiological basis for

M Barad
University of Los Angeles, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

General Considerations
Physiological Measures
Brain Imaging

Glossary

α_2-Adrenergic receptor	A receptor for the neurotransmitter norepinephrine; largely expressed presynaptically and acts to inhibit further release of norepinephrine.
Adrenergic system	The cells and pathways in the brain that use the neurotransmitter norepinephrine, either by releasing it or by expressing receptors that respond to it.
Corticotropin releasing hormone (CRH)	A peptide neurotransmitter released by the median eminence of the hypothalamus to stimulate the hypothalamic-pituitary-adrenal axis. It also acts widely in the brain as a neurotransmitter.
Cortisol	The primary glucocorticoid hormone.
Denditic pruning	Reductions in the branching dendrites of neurons, the part of the neuron that receives signals from other neurons.
Functional imaging	The examination of the brain by magnetic resonance imaging or positron emission tomography, yielding cross-sectional images of the brain and highlighting areas of increased or decreased blood flow that reflect localized brain activity.
Gluco-corticoids	Steroid hormones, especially cortisol, that regulate metabolism and the stress response.
Hippocampus	A brain structure crucially involved in complex memories (those available to consciousness).
Hypothalamic-pituitary-adrenal (HPA) axis	The regulatory pathway from the brain to the adrenal cortex that regulates the release of cortisol.
Longitudinal studies	Studies that follow the development of a disease in an individual over time, along with its biological correlates. By comparing variables before and after the development of a disease, these studies allow a stronger identification of the pathological effects of the disease process than other types of studies.
Metachlorophenylpiperazine (m-CPP)	A serotonin agonist.
Naltrexone	An antagonist of opioids receptors.
Opioid system	The cells and pathways in the brain that use endorphins, or endogenous opioids, as neurotransmitters, either by releasing them or expressing receptors that respond to them. These are the same receptors that bind opium and other opioid drugs.
Provocation studies	Studies that use drugs or environmental stimuli (e.g., stories, images, or tasks) to provoke a specific brain response.
Serotonin system	The cells and pathways in the brain that use the neurotransmitter serotonin, either by releasing it or by expressing receptors that respond to it.
Structural imaging	The examination of the brain using high-resolution magnetic resonance imaging (or, in the past, computed tomography) to carefully quantify the size of identifiable structures.
Twin studies	Research comparing identical twins, one of whom suffers a pathology, such as posttraumatic stress disorder. Because the healthy twin is genetically identically and shares much the same environment as the patient, the similarities in his

brain structure or function to that of the patient, which differ from the average population and are similar to the patient population, suggest a preexisting condition or vulnerability to disease.

Yohimbine A pharmacological antagonist of α_2-adrenergic receptor.

General Considerations

In general, two approaches have been used to examine the neurobiological basis of posttraumatic stress disorder (PTSD): imaging and biological measures. Imaging has been used to examine the anatomy of the brain and its activity both at rest and during a number of symptom provocation and cognitive challenges. Biological measures have been used to measure a number of candidate hormonal and neurotransmitter systems, again, both at rest and after symptom provocation. We consider some of the findings of these studies here.

It is important to remember that most studies of PTSD compare PTSD patients to normal controls. Such studies cannot differentiate between a finding in the PTSD patient that is the result of the disease process of PTSD and follows the traumatic event and a finding that preceded the trauma and thus may indicate a source of vulnerability to PTSD. Only longitudinal and twin or family studies can effectively differentiate between those possibilities, and there have been few such studies to date.

It is likely that preexisting factors, whether genetic or acquired from previous experience, are crucial to the development of PTSD. Exposure to trauma is not enough, no matter how severe. Only a small proportion of people exposed to trauma, even horrific trauma, develop PTSD. In the United States, approximately 90% of the population is exposed to at least one traumatic event serious enough to satisfy the trauma criterion for PTSD. However, only approximately 8% of the population develops PTSD. This means that it is particularly important to understand the neurobiological factors that contribute to vulnerability to PTSD or to its inverse, resiliency.

One obvious biological factor that seems to play a role in vulnerability to PTSD is sex. Women are twice as likely to develop PTSD after a trauma than men, although the interaction between biology and type of trauma is a complicated one. Women are much more likely to develop PTSD after an assault and men are much more likely to develop PTSD after being raped (even though their risk of rape is lower); both sexes have an equal risk of PTSD after natural disaster. These patterns suggest that the differences in rates of PTSD between men and women overall have something to do with the meaning of the type of trauma to each sex. Both sexes suffer equally from natural disasters, which affect everyone in the area, whereas physical or sexual assault emphasizes the relative weakness of the victim.

PTSD is, by definition, a chronic condition in which symptoms persist for at least 1 month. However, in the immediate aftermath of a trauma, almost everyone experiences symptoms of PTSD. Thus, the symptoms of PTSD are almost certainly part of the normal human response to stress. It is the failure of recovery that is characteristic of PTSD, so it may be that the right place to look for vulnerability is in the mechanisms of recovery. Such mechanisms may include biological mechanisms, such as the glucocorticoid response to stress, or psychological mechanisms, such as the ability of people to suppress fearful responding in the short term through inhibition by prefrontal cortex or in the long term through behavioral extinction or desensitization of fearful responses to the reminders of the trauma.

Physiological Measures

Frightening experiences elicit a characteristic stress response. The faster response to a frightening stimulus is the adrenergic stress response, which orchestrates the readiness of an animal or a human to protect itself, the fight-or-flight response originally described by Walter Cannon. Simultaneously, a second, slower stress system is activated, the hypothalamic-pituitary-adrenal (HPA) axis, which results in the production of the effector glucocorticoid cortisol, as first described by Hans Selye. One important action of the slower cortisol response is to shut down or limit the adrenergic response.

Disruptions in stress hormone regulation stand out prominently in the neurobiology of PTSD. Some of the first studies of PTSD patients indicated not only that PTSD patients show increased excretion of epinephrine and norepinephrine at baseline but also that adrenergic system is hyperreactive. The administration of yohimbine to PTSD patients provided the most striking demonstration of the increase of adrenergic tone and its relation to PTSD symptoms. Yohimbine blocks the α_2 receptors for norepinephrine. These receptors are presynaptic on adrenergic terminals and act like a thermostat to inhibit the release of norepinephrine into the synapse. The administration of yohimbine is mildly anxiogenic in normal controls, but in PTSD patients it caused great distress, precipitating flashbacks in 40% and panic attacks in 70%, whereas none of the controls experienced either. These data indicate a profound dysregulation of the adrenergic system. Interestingly, yohimbine

also provokes panic attacks in patients with panic disorder but not in patients with schizophrenia, depression, obsessive-compulsive disorder, or generalized anxiety disorder and also not in normals. Thus, PTSD may share some part of its pathophysiology with panic disorder.

Increases of adrenergic tone and reactivity may be of great importance in the symptomatology of PTSD. It is clear from the experiment that increases of adrenergic neurotransmission may be precipitating some of the most troubling symptoms of PTSD, including not only flashbacks and the physiological hyperreactivity characteristic of panic attacks but also intrusive thoughts, emotional numbing, difficulty concentrating, and increased startle. However, it has also been extensively demonstrated in preclinical studies that the adrenergic system is crucial to memory strength, and particularly to the establishment of fearful or distressing memories. This explains in part why memories of frightening events are so vivid. Thus, high levels of norepinephrine during traumatic experiences may well strengthen the formation of distressing memories and make them more vivid for PTSD patients.

When challenged with yohimbine, the family members of PTSD patients do not have more panic attacks than normal subjects, whereas the healthy family members of panic disorder patients do have high rates of panic attacks upon yohimbine challenge. This suggests that a hyperreactive adrenergic system may be a predisposing factor for panic disorder, whereas the adrenergic hyperreactivity of PTSD patients may be acquired.

What might account for that hyperreactivity? One hypothesis is an insufficient regulation by an inadequate glucocorticoid response. We might expect from the strong response to trauma in PTSD patients that cortisol levels, as part of the stress response, should be high. In fact, most studies have found that glucocorticoid levels are low at baseline in PTSD patients and that the feedback inhibition of glucorticoid synthesis is overly strong. Thus, when injected with an exogenous glucocorticoid, dexamethasone, patients with PTSD suppress the production of endogenous cortisol for a significantly longer time than normal controls. This is true even though the levels of the hypothalamic hormone that drives cortisol synthesis and release, corticotropin releasing hormone (CRH), are high in PTSD patients.

By contrast, in depression, high CRH levels drive high levels of circulating cortisol. The contrasting pattern in PTSD suggests that feedback inhibition of the HPA axis is dominant over CRH drive. Low circulating levels of cortisol may well allow the adrenergic response to stress to run out of control. But are these low levels in PTSD patients the result of a burnout of the cortisol system by sustained high cortisol during and after the trauma?

Two longitudinal studies of motor vehicle-accident victims have addressed this question. In one study, 15-h urinary cortisol excretion was measured after accidents; in the other, cortisol levels were measured in blood samples drawn on the way to the emergency room. In both studies, the victims who went on to develop PTSD had significantly lower cortisol after their accidents than those who did not. Thus, even just after their trauma, the cortisol levels of those who later went on to PTSD were lower. This begins to hint that lower cortisol in PTSD patients may not be the result of burnout after prolonged elevations around the time of trauma but rather a preexisting state that contributes to a vulnerability to PTSD.

This preexisting state is almost certainly not entirely genetic. For example, in rape victims, cortisol levels after the rape were lower in those who had sustained a previous assault, and the victims of previous assaults were three times as likely to go on to develop PTSD. Women who had no previous trauma had cortisol levels that averaged three times higher than those who did and had higher cortisol levels when the index rape was more severe. Women with previous trauma showed no increase in cortisol levels in response to increased severity of rape.

However, the question remains whether low cortisol levels existed even before the rape. By comparing the children of Holocaust survivors to children of European Jews who escaped the Holocaust, Rachel Yehuda and her colleagues have examined this question. The children of holocaust survivors are almost four times as likely to develop PTSD as the controls. The children of Holocaust survivors show lower cortisol excretion than do the controls, which argues for a preexisting defect in the cortisol system. However, this group was not homogeneous. When neither parent nor child had PTSD, cortisol excretion was the same as that of controls. However, when the parent had PTSD, the cortisol excretion was lower in the offspring; it was lowest when both parent and child had PTSD. Thus, it is clear that low PTSD levels are a preexisting risk factor that may not only be due to previous trauma. It is tempting to think that this risk factor is due to the transmitted anxiety of parents with PTSD, acting like a previous assault in rape victims to lower cortisol. However, there remains a potential for a genetic explanation. It might be that the Holocaust survivors with PTSD also had low cortisol as a risk factor for their PTSD and transmitted that risk genetically rather than behaviorally.

All of these data on endocrine stress responses outline one potential chain of causation in generating

the clinical picture of PTSD. A hypoactive cortisol response may lead to the dysregulation of the adrenergic stress response. An ungoverned adrenergic system may increase the intensity of traumatic memories, the strength of association of those memories with a wide range of stimuli in the environment, and the physiological and psychological arousal when those memories are recalled.

Beyond the alterations in the endocrine stress responses, other physiological abnormalities have been identified in PTSD patients. Endogenous opioids are low in PTSD sufferers, who also show a reduced pain threshold. However, when exposed to traumatic reminders, Vietnam veterans with PTSD showed a higher pain threshold that could be reversed by naltrexone (which blocks opioids receptors). Thus, PTSD patients appear to have both decreased resting levels of endogenous opioids and a more reactive opioid system. This may lead to cycles of trauma reminders inducing numbing, followed by opioid withdrawal (and PTSD) symptoms such as irritability and disturbed sleep. Because these hyperarousal symptoms of opioid withdrawal may be mediated by adrenergic activity, low tonic opioid levels may act in concert with low cortisol to increase adrenergic reactivity.

One study examined the effect of a serotonin agonist metachlorophenylpiperazine (m-CPP) on PTSD patients compared with the effect of yohimbine and placebo. As in the studies of yohimbine, m-CPP caused panic attacks in a substantial number of the patients, along with exacerbations of other PTSD symptoms. A similar number of patients experienced panic attacks and worsened PTSD symptoms after m-CPP injections. Interestingly, different subsets of patients tended to get panic attacks with the two agents. Another study showed that PTSD patients had a higher frequency of being homozygous for the short allele of the serotonin promoter, a genotype associated with depression after life stressors. Thus, although fewer data indicate a role for serotonergic than for noradrenergic mechanisms, serotonergic mechanisms may make a second and possibly independent contribution to PTSD vulnerability or pathophysiology.

Brain Imaging

Three areas of the brain have received particular attention in imaging studies of PTSD, because there are *a priori* reasons to suspect their involvement. The hippocampus has been examined in detail because a wealth of preclinical studies indicate that chronic stress has profound effects on this structure. The amygdala has been examined because this brain structure has an orchestrating role in all fear learning. Finally, the prefrontal cortex has been examined because of its role in suppressing amygdaloid activation.

Structural Imaging

A variety of studies in animals indicate that elevated levels of the stress hormone cortisol is damaging to the hippocampus. Long-term exposure to exogenous glucocorticoid steroids and prolonged stress (which presumably elevates endogenous levels of these steroids) both damage the CA3 region of the hippocampus and cause dendritic pruning of neurons there, with the net effect of decreasing hippocampal volume. Although lower levels of cortisol are usually observed in PTSD patients, it is reasonable to hypothesize that cortisol levels might have been extremely high at the time of traumatic exposure.

Consistent with this hypothesis, hippocampal volumes in PTSD patients are decreased compared to controls in a number of structural imaging studies of combat veterans and of adult victims of childhood sexual abuse that compared PTSD patients to others exposed to combat or sexual abuse without PTSD. However, results have been less consistent with other patient populations and have not been observed in Holocaust survivors or patients with personality disorders who suffer from PTSD.

These studies are cross-sectional, comparing PTSD patients to controls, and do not address whether the hippocampal volume deficits might predate the trauma and development of PTSD. One study of identical twins in which one twin was a Vietnam combat veteran addressed this question indirectly. This study found that smaller hippocampal volume in the combat-exposed twin correlated to the severity of his PTSD symptoms. However, the hippocampal volume of the non-combat-exposed twin correlated just as closely with the severity of his brother's symptoms. These data suggest strongly that small hippocampi are a risk factor for, rather than a result of, combat exposure. Two longitudinal studies have also addressed this question, and neither found decreases of hippocampal volume associated with the development of PTSD or with ongoing PTSD pathology. Similarly, no changes in amygdala volume have been seen in PTSD patients.

Functional Imaging

Brain structure may be of less importance than brain function in psychiatric disease, including in PTSD. Functional imaging, looking at brain activity as reflected by blood flow, may thus be a better indicator of pathological brain function, especially when performed using biological or psychological probes designed to stimulate functional systems. Functional

studies have focused attention on the amygdala, a brain region important for orchestrating the fear response, as well as in related paralimbic and frontal structures. This approach remains in its early stages, and a clear anatomical pattern has not yet emerged, perhaps because of the wide variation in patient populations and provocation protocols in use.

One interesting approach has been to use pharmacological provocation with yohimbine, the antagonist of inhibitory presynaptic α_2 autoreceptors. As previously described, yohimbine, when injected into PTSD patients, exascerbates symptoms and causes panic attacks and flashbacks. In a positron emission tomography (PET) imaging study, yohimbine injections tended to increase metabolism in a variety of cortical regions including the prefrontal, temporal, parietal, and orbitofrontal cortices in control subjects, but tended to decrease metabolism in those regions in PTSD patients. These data indicate that the adrenergic dysregulation observed in PTSD patients profoundly changes brain activation patterns and suggests a role for those brain structures with changed activity in suppressing the set of symptoms evoked by yohimbine, specifically panic attacks and flashbacks.

Psychologically driven provocation studies have generally shown differences in brain activation between PTSD patients and controls, although these have not always been consistent. Left amygdala activity was increased by combat sounds, compared to white noise, in Vietnam veterans with PTSD but not in combat-exposed veterans without PTSD or in normal controls. In another similar study with combat sound, veterans with PTSD showed activation of the right amygdala as well as in other regions. In another comparison of veterans with and without PTSD who were asked to visualize combat, the veterans showed activation in the right amygdala as well as in the ventral anterior cingulate gyrus. However, when actually viewing combat images, no such activation was seen; however, PTSD patients showed decreased activity in Broca's area, an area important for language generation.

Trauma scripts and viewing of traumatic images have generally not yielded amygdala activation, whereas visualization and simple stimuli, such as combat sounds, which bypass the cortex, effectively activate that structure. Interestingly, masked presentations of fearful faces, too brief to register consciously, caused exaggerated amygdala responses in veterans with PTSD compared to healthy veterans.

On the other hand, script-driven provocation has effectively activated the posterior cingulate and motor cortex in women with a history of childhood sexual abuse with PTSD compared to those without PTSD. Women with PTSD also failed to activate Broca's area and the anterior cingulate compared to controls, who did.

In summary, the functional imaging data so far do support changes in the function of the amygdala and of the limbic and prefrontal cortices in PTSD. However, the exact nature of those changes and their significance in the symptomatology and pathophysiology of PTSD remain to be defined.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Catecholamines; Corticotropin Releasing Factor (CRF); Corticotropin-Releasing Factor Receptors; Endocrine Systems; Fight-or-Flight Response; Glucocorticoid Negative Feedback; Glucocorticoids, Effects of Stress on; Glucocorticoids – Adverse Effects on the Nervous System; Glucocorticoids, Overview; Glucocorticoids, Role in Stress; Hippocampus, Corticosteroid Effects on; Hypothalamic-Pituitary-Adrenal; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of; Posttraumatic Therapy; Sex Steroids, Response to Stress and Susceptibility to Depression; Fear and the Amygdala; Neuroimaging and Emotion; Opioids; Selye, Hans; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Clinical; Serotonin in Stress.

Further Reading

Bremner, J. D. (2002). Neuroimaging studies in posttraumatic stress disorder. *Current Psychiatry Reports* **4**, 254–263.

Bremner, J. D. (2005). Effects of traumatic stress on brain structure and function: relevance to early responses to trauma. *Journal of Trauma and Dissociation* **6**, 51–68.

Foa, E. B. and Street, G. P. (2001). Women and traumatic events. *Journal of Clinical Psychiatry* **62**(supplement 17), 29–34.

Gilbertson, M. W., Shenton, M. E., Ciszewski, A., et al. (2002). Smaller hippocampal volume predicts pathologic vulnerability to psychological trauma. *Nature Neuroscience* **5**, 1242–1247.

Grossman, R., Buchsbaum, M. S. and Yehuda, R. (2002). Neuroimaging studies in post-traumatic stress disorder. *Psychiatric Clinics of North America* **25**, 317–340.

Resnick, H. S., Yehuda, R., Pitman, R. K., et al. (1995). Effect of previous trauma on acute plasma cortisol level following rape. *American Journal of Psychiatry* **152**, 1675–1677.

Rinne, T., Westenberg, H. G., den Boer, J. A., et al. (2000). Serotonergic blunting to meta-chlorophenylpiperazine (m-CPP) highly correlates with sustained childhood

abuse in impulsive and autoaggressive female borderline patients. *Biological Psychiatry* **47**, 548–556.
Southwick, S. M., Bremner, J. D., Rasmusson, A., et al. (1999). Role of norepinephrine in the pathophysiology and treatment of posttraumatic stress disorder. *Biological Psychiatry* **46**, 1192–1204.
Vermetten, E. and Bremner, J. D. (2002). Circuits and systems in stress. II: Applications to neurobiology and treatment in posttraumatic stress disorder. *Depression and Anxiety* **16**, 14–38.
Yehuda, R. (2002). Current status of cortisol findings in post-traumatic stress disorder. *Psychiatric Clinics of North America* **25**, 341–368.
Yehuda, R. (2004). Risk and resilience in posttraumatic stress disorder. *Journal of Clinical Psychiatry* **65**(supplement 1), 29–36.
Yehuda, R., Schmeidler, J., Wainberg, M., et al. (1998). Vulnerability to posttraumatic stress disorder in adult offspring of Holocaust survivors. *American Journal of Psychiatry* **155**, 1163–1171.

Posttraumatic Stress Disorder in Children

A M La Greca
University of Miami, Coral Gables, FL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A M La Greca, volume 3, pp 181–185, © 2000, Elsevier Inc.

Traumatic Events
Posttraumatic Stress Disorder in Children: Clinical Picture
Prevalence and Developmental Course
Demographic Trends
Factors that Precipitate Posttraumatic Stress Disorder in Children
Assessing Posttraumatic Stress Disorder Symptoms in Children

Glossary

Avoidance or numbing	Refers to symptoms of posttraumatic stress syndrome (PTSD), such as avoiding thoughts, feelings, or conversations about the traumatic event, avoiding reminders of the event, diminished interest in normal activities, or feeling detached or removed from others.
Hyperarousal	Refers to symptoms of PTSD, such as difficulty sleeping or concentrating, irritability, angry outbursts, hypervigilance, and an exaggerated startle response.
Life threat	The perception that one's life is in danger; it is a key aspect of exposure to traumatic events, and is believed to be critical for the emergence of PTSD and related symptoms.
Loss and disruption	Loss (of family and friends, of personal property and possessions) and disruption of everyday life (displacement from home, school, or community); it is also an aspect of exposure to traumatic events and can contribute to PTSD in children.
Re-experiencing	Refers to symptoms of PTSD, such as recurrent or intrusive thoughts or dreams about the event, or intense distress at cues or reminders of the event.

Posttraumatic stress disorder (PTSD) refers to a set of symptoms that develop following exposure to an unusually severe stressor or event. Typically, the event is one that causes or is capable of causing death, injury, or threat to the physical integrity of oneself or another person. In order to meet criteria for a diagnosis of PTSD, a child's reaction to the traumatic event must include intense fear, helplessness, or disorganized behavior. In addition, specific criteria for three additional symptom clusters must be met: re-experiencing, avoidance/numbing, and hyperarousal. Re-experiencing includes symptoms such as recurrent or intrusive thoughts or dreams about the event and intense distress at cues or reminders of the event; for young children, re-experiencing also may be reflected in repetitive play with traumatic themes or by a re-enactment of traumatic events in play, drawings, or verbalizations. Avoidance or numbing includes symptoms such as efforts to avoid thoughts, feelings, or conversations about the traumatic event, avoiding reminders of the event, diminished interest in normal activities, and feeling detached or removed from others. Hyperarousal symptoms include difficulty

sleeping or concentrating, irritability, angry outbursts, hypervigilance, and an exaggerated startle response; these behaviors must be newly occurring since the event. For a diagnosis of PTSD, the above symptoms must be manifest for at least 1 month, and be accompanied by significant impairment in the child's functioning (e.g., problems in school, friendships, or family relations). In addition, acute PTSD is specified if the duration of the symptoms is less than 3 months; chronic PTSD is specified if the duration of symptoms is 3 months or longer.

Traumatic Events

In recent years, media coverage of devastating natural disasters, bombings, and terrorist activities has drawn attention to the significant trauma that children and adults experience following such events. It has become increasingly apparent that exposure of children to such traumatic events may lead to reactions that interfere substantially with their day-to-day functioning and cause them significant distress. Specifically, exposure to natural disasters (e.g., hurricanes, tsunamis, tornadoes, fires, earthquakes, or floods) and to human-made disasters and acts of violence (e.g., plane crashes, sinking of ferrys, war, terrorist attacks, bombings, sniper shootings) represent traumatic events that may result in the emergence of PTSD. Although considerable attention has been devoted to the effects of terrorism and community-wide disasters, it is also the case that children who experience motor vehicle accidents, life-threatening illnesses or medical procedures, or violence of a personal nature (e.g., physical and sexual abuse, rape, kidnapping, and community violence) also are vulnerable to developing PTSD. This article describes common symptoms of PTSD in children and factors that contribute to the development and course of this disorder.

Posttraumatic Stress Disorder in Children: Clinical Picture

The diagnostic category of PTSD was introduced in the third edition of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III). At that time, PTSD was primarily considered to be an adult disorder. More recently, there has been a growing awareness that children and adolescents also experience PTSD, which is reflected in the current edition of the DSM (DSM-IV). (See previous description of the criteria for the disorder.)

Community studies of children following traumatic events suggest that the most commonly reported PTSD symptoms are those of re-experiencing the event. For example, up to 90% of children exposed to a catastrophic hurricane reported symptoms of re-experiencing 3 months after the disaster. In contrast, symptoms of avoidance and numbing are far less common in children and, in fact, their presence may be a good indicator of PTSD.

Due to their limited verbal capacity, the diagnosis of PTSD is especially difficult in very young children, such as infants, toddlers, and early preschoolers. In such cases, generalized anxiety, fears, avoidance of certain situations that may be linked to the traumatic event, and sleep disturbances may be useful indicators of PTSD.

PTSD may also be present (i.e., co-morbid) with other psychological difficulties. In particular, PTSD often is co-morbid with anxiety disorders (i.e., separation anxiety disorder, specific phobias, generalized anxiety disorder, panic, and agoraphobia) and with affective disorders (e.g., major depressive disorder and dysthymia). Traumatic events that involve the loss of loved ones (e.g., earthquakes and bombings) are especially likely to lead to depressive reactions in children.

Prevalence and Developmental Course

Prevalence

It is difficult to estimate the prevalence of PTSD in children and adolescents because studies have been extremely diverse with respect to the type of trauma evaluated, assessment methods used, and the length of time that passed since the traumatic event occurred. Community studies suggest that approximately 24–39% of children and adolescents exposed to trauma (e.g., community violence or a natural disaster) meet criteria for PTSD. Rates of PTSD have been reported to be even higher among children and adolescents who witness death or physical injury in conjunction with acts of violence or following disasters associated with mass casualties.

When subclinical levels of PTSD are considered, up to 55% of the children in large community samples have reported at least moderate levels of PTSD symptoms during the first 3 months following a traumatic event, such as a hurricane. Thus, symptoms of PTSD appear to be common among children exposed to trauma, although fewer children and youth will meet criteria for a PTSD diagnosis.

Developmental Course

Little is known about the course of PTSD symptoms in children over time. However, it does appear that PTSD symptoms may emerge in the days or weeks

following a traumatic event, and can take months or years to dissipate in some children. In the absence of re-exposure to trauma or of the occurrence of other traumatic events, the typical developmental course of symptoms appears to be one of lessening frequency and intensity over time. For example, 3 months after a highly destructive natural disaster (Hurricane Andrew: a Category 5 hurricane, that struck South Florida in 1992), 39% of the children informally met criteria for PTSD disorder, but this was reduced to 24% at 7-months postdisaster, and to 18% by 10-months postdisaster. A subgroup of children who reported moderate-to-severe PTSD symptoms was followed 42 months postdisaster, revealing that 40% continued to report moderate-to-severe PTSD symptoms as well as impairment in their functioning; yet, almost none of the children who had mild or no symptoms at 10-months postdisaster reported any symptoms later on. Similarly, 51.5% of the adolescents who survived the sinking of the cruise ship Jupiter in 1988 (with 400 British school children on board) developed PTSD after the disaster, but 5–8 years later, only 17.5% still had PTSD.

These data suggest a steady reduction in the frequency and severity of PTSD and its symptoms over time (with no further exposure to similar trauma), although a significant minority of children and adolescents do not recover and continue to report substantial difficulties years later. The findings also indicate that it is highly unusual for children and adolescents to report significant symptoms of PTSD 1 year or more after a traumatic event, if they did not experience such symptoms closer to the event. Although there may be a brief period of shock or numbing, or sometimes elation and relief at still being alive, most children who develop PTSD will evidence stress symptoms within the first few months of the traumatic event.

Persistence of Symptoms

Overall, findings confirm that children's reactions to disasters and other traumatic events are not merely transitory feelings that dissipate quickly. On the contrary, they appear to linger and persist and are likely to cause distress to children and their families for some time. One factor that contributes to the persistence of PTSD symptoms in children over time is the occurrence of other significant life stressors. Children who encounter major life stressors (e.g., a death or illness in the family or parental divorce) in the months following a traumatic event, do not recover quickly, and report persistently high levels of PTSD symptoms over time.

Demographic Trends

Gender and Age

Some studies indicate that PTSD appears more frequently among girls than boys, but this has not consistently been the case. Preadolescent children may be more vulnerable to PTSD than older youths, although it is difficult to draw generalizations regarding children's vulnerability to PTSD at different ages, as findings on age-related differences have been inconsistent and may be influenced by diverse developmental manifestations of PTSD.

Ethnicity

PTSD appears to occur across diverse ethnic and cultural groups. Community studies suggests that minority youths exposed to severe natural disasters report more symptoms of PTSD and have a more difficult time recovering from such events than nonminority youths. It is possible that socioeconomic factors might account for such findings, in that children and families from minority backgrounds may have less financial resources or less adequate insurance to deal effectively and efficiently with the postdisaster rebuilding process. This could prolong the period of life disruption and loss of possessions that ensues after destructive natural disasters.

Factors that Precipitate Posttraumatic Stress Disorder in Children

Exposure to Trauma

A wide range of traumatic events have been linked to the emergence of PTSD in children and adolescents. However, two key aspects of exposure to these traumatic events are thought to contribute to children's reactions: the presence or perception of life threat, and personal loss and disruption of everyday events.

The presence or perception of life threat is considered to be essential for the emergence of PTSD. It is easy to understand why children who witness or are exposed to acts of violence, such as sniper shootings or severe physical abuse by a caretaker, feel that their life is in danger. It is also the case, however, that catastrophic natural disasters or personal disasters (such as residential fires) can elicit perceptions of life threat in children, even if no one is injured or hurt. For example, although very few lives were lost in South Florida as a result of Hurricane Andrew, the extensive destruction of homes and property that occurred during the storm was terrifying to many children and adults. In one study, 60% of the children interviewed "thought that they were going to die"

during the storm. Thus, perceptions of life threat can occur in the absence of actual loss of life or serious injury.

Evidence also underscores the importance of loss (of family and friends or personal property and possessions) and disruption of everyday life (displacement from home, school, or community) as contributing to PTSD in children. The life changes that result from this aspect of exposure to a traumatic event also predicts PTSD symptoms in children and adolescents.

The duration and intensity of life-threatening events are additional aspects of traumatic exposure associated with children's PTSD symptom severity. The prolonged nature of certain traumatic events (e.g., floods or child abuse) is very distressing to children when no immediate relief is in sight.

Role of Prior Psychological Functioning

Children with pre-existing psychological problems, especially anxiety, may be more vulnerable to developing PTSD following traumatic events. For example, children's anxiety levels 15 months before a traumatic event have been found to predict their levels of PTSD symptoms 3 and 7 months postdisaster, even when controlling for children's exposure to the event. This suggests that anxious children may have a vulnerability to PTSD, even if their degree of exposure to a trauma is relatively low. In addition, children who had greater levels of exposure to the disaster (i.e., more life threat or more loss/disruption) showed an increase in anxiety symptoms following the disaster relative to children with less disaster exposure. Other studies have shown predisaster levels of depression predicted postdisaster PTSD among adolescents exposed to a destructive natural disaster (earthquake).

Aspects of the Recovery Environment

Certain aspects of the posttrauma recovery environment may either contribute to the development and persistent of symptoms of PTSD, or may help to alleviate or mitigate such symptoms. Intervening life events/stressors and re-exposure to the traumatic event have been linked with greater persistence of PTSD and its symptoms among children and adolescents. In addition, parents' psychosocial functioning, including their levels of psychopathology and their own reactions to the disaster, are likely to affect children's postdisaster functioning. For example, mothers' distress in the aftermath of Hurricane Hugo, which struck Charleston South Carolina in 1989, was associated with the persistence of their children's postdisaster emotional difficulties. Other research has supported a linkage between children's symptoms of PTSD and parents' trauma-related symptoms. Finally, children with more negative coping strategies for dealing with stress (e.g., anger or blaming others) show greater persistence in their symptoms of PTSD over time.

In contrast, the availability of social support has been found to mitigate the impact of trauma on children. Following traumatic events, such as severe abuse or catastrophic disasters, children with higher levels of social support from significant others (parents, teachers, friends, and classmates) report fewer symptoms of PTSD than those with low levels of social support. Because of these findings, efforts to enhance the social support of children exposed to trauma, and to encourage adaptive coping skills, may be useful for strategies for interventions with children following traumatic events.

Assessing Posttraumatic Stress Disorder Symptoms in Children

Evaluating the presence of PTSD symptoms in children and adolescents can be challenging and will depend, to some extent, on the age or developmental level of the child. For children of elementary school age and older (i.e., 7 years and above), the child or adolescent is likely to be the best informant; parents often underestimate symptoms of PTSD in their children.

The use of child-oriented structured interviews that contain items pertinent to the PTSD diagnosis, such as the Anxiety Disorders Interview Schedule for Children or the Diagnostic Interview Schedule for Children represent the most desirable way to assess such reactions. For screening purposes, when a large number of children are being evaluated, or when personnel constraints preclude the use of individual interviews, child-oriented self-report measures, such as the PTSD Reaction Index, have been found to be extremely useful.

Assessing symptoms of PTSD in children of preschool age and younger typically requires the input of the parent or primary carer. Parent-versions of structured interviews would be the desired format to use, or parent-competed questionnaires reporting on children's PTSD symptoms. For young children, behavioral observations can be especially important in evaluating their reactions. Behaviors to evaluate include signs of distress, arousal, or fear, and behavioral avoidance of trauma-related objects, events, or situations.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Posttraumatic Stress Disorder, Delayed;

Posttraumatic Stress Disorder, Neurobiology of; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder – Neurobiological basis for.

Further Reading

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). (DSM-IV) Washington, DC: American Psychiatric Association.

American Academy of Child Adolescent Psychiatry (1998). AACAP Official Action: Practice parameters for the assessment and treatment of children and adolescents with posttraumatic stress disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* **37**, 4S–26S.

Applied Research & Consulting LLC and the Columbia University Mailman School of Public Health. (2002). *Effects of the World Trade Center attack on NYC public school students: initial report to the New York City Board of Education*. Available online at:. http://eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/27/ce/00.pdf. Last accessed 11 October 2006.

Asarnow, J. (1999). When the earth stops shaking: earthquake sequelae among children diagnosed for pre-earthquake psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry* **38**, 1016–1023.

Drell, M. J., Siegel, C. H. and Gaensbauer, T. J. (1993). Posttraumatic stress disorder. In: Zeanah, C. H. (ed.) *Handbook of infant mental health*, pp. 369–381. New York: Guilford Press.

Goenjian, A. K., Pynoos, R. S., Steinberg, A. M., et al. (1995). Psychiatric comorbidity in children after the 1988 earthquake in Armenia. *Journal of the American Academy of Child and Adolescent Psychiatry* **34**, 1174–1184.

Green, B. L., Korol, M. S., Grace, M. C., et al. (1991). Children and disaster: gender and parental effects on PTSD symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry* **30**, 945–951.

Gurwitch, R. H., Sitterle, K. A., Young, B. H., et al. (2002). The aftermath of terrorism. In: La Greca, A. M., Silverman, W. K., Vernberg, E. M., et al. (eds.) *Helping children cope with disasters and terrorism*, pp. 327–358. Washington, DC: American Psychological Association.

La Greca, A. M., Silverman, W. K., Vernberg, E. M., et al. (1996). Symptoms of posttraumatic stress after Hurricane Andrew: a prospective study. *Journal of Consulting and Clinical Psychology* **64**, 712–723.

La Greca, A. M., Silverman, W. K., Vernberg, E. M., et al. (eds.) *Helping children cope with disasters and terrorism*. Washington, DC: American Psychological Association.

La Greca, A. M., Silverman, W. K. and Wasserstein, S. B. (1998). Children's predisaster functioning as a predictor of posttraumatic stress following Hurricane Andrew. *Journal of Consulting and Clinical Psychology* **66**, 883–892.

Nolen-Hoeksema, S. and Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: the 1989 Loma Prieta Earthquake. *Journal of Personality and Social Psychology* **61**, 115–121.

Vernberg, E. M., La Greca, A. M., Silverman, W. K., et al. (1996). Predictors of children's post-disaster functioning following Hurricane Andrew. *Journal of Abnormal Psychology* **105**, 237–248.

Vogel, J. and Vernberg, E. M. (1993). Children's psychological responses to disaster. *Journal of Clinical Child Psychology* **22**, 464–484.

Yule, W., Udwin, O. and Bolton, D. (2002). Mass transportation disasters. In: La Greca, A. M., Silverman, W. K., Vernberg, E. M., et al. (eds.) *Helping children cope with disasters and terrorism*, pp. 327–358. Washington, DC: American Psychological Association.

Posttraumatic Stress Disorder Overview, General *See:* Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder, Neurobiology of; Posttraumatic Therapy; Acute Stress Disorder and Posttraumatic Stress Disorder.

Posttraumatic Stress Disorder, Delayed

A Holen
Norwegian University of Science and Technology, Trondheim, Norway

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Holen, volume 3, pp 179–180, © 2000, Elsevier Inc.

Diagnostic Systems and Delayed Posttraumatic Stress Disorder
Nature of the Delay Period
Discussing Delayed Posttraumatic Stress Disorder

Glossary

Delayed posttraumatic stress disorder (D-PTSD)	Posttraumatic stress disorder when the onset of symptoms occurs at least 6 months after the stressor.
Stressor	A traumatic event that gives rise to major stress responses.

Diagnostic Systems and Delayed Posttraumatic Stress Disorder

Posttraumatic stress disorder (PTSD) can be found in both the *Diagnostic and Statistical Manual of Mental Disorders,* 4th edition (DSM-IV) of the American Psychiatric Association and the *International Classification of Diseases* (ICD-10) of the World Health Organization. The stressor criterion and the three symptom cluster criteria (intrusion, avoidance, and psychophysiological arousal), as well as the time frames, are fairly similar, though some differences exist.

In DSM-III through DSM-IV, PTSD is subsumed under anxiety disorders. When the disorder lasts at least 1 month and persists for less than 3 months after the stressful event, PTSD is considered acute. If the duration is more than 3 months, then PTSD is regarded as chronic.

In ICD-10, PTSD is placed under the heading F43, "Reaction to severe stress and adjustment disorders." The ICD-10 defines "Post-traumatic stress disorder" (F43.1) when the diagnostic criteria are met within 6 months of a major stressful event or of a period of stress. The ICD-10 gives yet another option: when patients have a chronic course, which involves major and lasting personality changes, the diagnosis of "Enduring personality changes, not attributable to brain damage and disease" (F62) may be used. So far, there has been limited research into long-term personality changes.

Both DSM-IV and ICD-10 concur that delayed PTSD (D-PTSD) is diagnosed when at least 6 months have passed between the exposure of a major stressor and the full onset of PTSD.

Nature of the Delay Period

D-PTSD occurs when a person is reasonably well adapted for the initial 6 months or more after experiencing a major stressor but develops PTSD symptoms, such as intrusion, avoidance, and increased arousal related to the initial event, at some later stage. None of the diagnostic systems, however, characterize the nature of the delay period. Still, there is a limited body of research available to shed light on the subject. North et al., in a study published in 1997, found no cases of delayed-onset PTSD in 136 survivors 1 year after a mass shooting incident. Considerable discrepancy of identified PTSD cases was apparent between index and follow-up. The authors argued that inconsistency in reporting rather than true delayed onset was responsible for PTSD cases identified 1 year after the trauma. After the Oklahoma City bombing, survivors were studied 6 and 17 months postdisaster by the same author. Again, no delayed-onset cases were verified. At a 5-year follow-up, Holen found only one case of D-PTSD in the 72 survivors from the North Sea oil rig disaster.

A prospective study by Bryant et al. of 103 motor vehicle accident survivors studied within 1 month of the stressor and again at 6 months and 2 years post-accident identified five patients with PTSD at 2 years; they had not met the PTSD criteria at 6 months. Additionally, the authors reported that delayed-onset cases suffered subsyndromal levels of PTSD, with elevated resting heart rate levels prior to diagnosis. Buckley et al. also found delayed-onset survivors from motor vehicle accidents to be more symptomatic initially than those who did not develop PTSD. In addition, the delayed-onset survivors tended to have poorer social support systems than controls both prior to and after the stressor. Furthermore, the delayed-onset survivors had lower global assessment of function (GAF) scores than controls in the month prior to the motor vehicle accident. Schnurr et al. found that women were more likely to be in the delayed-onset cluster. In a prospective study of traffic accident victims by Koren et al., only 2 of 39 survivors without PTSD at 1 year posttrauma demonstrated delayed-onset PTSD 4 years later. In discussing

compensation issues, Eitinger and Holen argued that bridge symptoms, i.e., subclinical symptoms that persist from the time of the trauma to the manifestations of full-blown PTSD, may serve as precursors to delayed responses. In follow-up studies of the population from the collapse of the Buffalo Creek dam, Green et al. reported that 11% of the cases not identified with PTSD in 1974 met the criteria in 1986.

In a study published in 1989, Solomon et al. reviewed 150 randomly selected files of veterans who sought treatment for combat-related disorders; all the veterans were from the 1982 Lebanon war. Four reasons were given for why they asked for health services. In 10% of the sample, the delayed-onset PTSD was preceded by an initial asymptomatic period. After the war, some other veterans uninterruptedly suffered mild PTSD-associated symptoms until accumulated tension or exposure to subsequent adversity, either military or civilian in nature, resulted in full-blown PTSD. Such exacerbations of PTSD were seen in 33% of the veterans. Reactivation of an earlier combat stress reaction was found in 13% of the cases. The reactivation started only after an asymptomatic period and in connection with threatening military stimuli. Some veterans had no single trigger; they seemed to build up tensions over time with various exposures. Finally, one group of veterans was characterized by delayed treatment seeking for chronic PTSD; this accounted for as much as 40% of the cases. In the absence of proper definitions, the first three groups, i.e., those with an initial asymptomatic period, those with exacerbations, and those with reactivated responses, may qualify as varieties of D-PTSD or they may represent different kinds of posttraumatic processing. The veterans from the 1982 Lebanon war with delayed-onset PTSD used significantly more emotion-focused and distraction coping on the Ways of Coping Checklist than control subjects. Furthermore, they used coping skills significantly less than those with immediate-onset PTSD.

From 1968 to 1973, decades before PTSD was invented, Eitinger and Stroem studied the health of Norwegian concentration camp survivors. They reported a high occurrence of late-onset somatic and psychiatric manifestations associated with concentration camp syndrome, or the KZ syndrome.

In a study of aging Resistance veterans from Holland, Aarts et al. found that the period between the end of the war and when the first symptoms of PTSD appeared varied considerably. First onset of PTSD symptoms appeared in 26% of the cases during the first 5 years after World War II. However, in about 50% of the veterans, PTSD manifested after more than 20 years.

Discussing Delayed Posttraumatic Stress Disorder

There are many uncertainties and controversies in the current conceptualizations of D-PTSD. Some authors report no cases of D-PTSD and raise doubts about the legitimacy of the concept. Others bring up the question of whether D-PTSD represents a latent disorder or simply is unrecognized PTSD. In recent years, a growing number of studies involving both motor vehicle accidents and combat-related stressors have demonstrated occurrence rates of D-PTSD around 10% when manifestation of the disorder is preceded by an asymptomatic period. Moreover, several studies point toward vulnerability factors in both the survivors and their environments. Persons who tend to have subsyndromal symptom levels in the delay period are more likely to have poorer social networks, poorer coping skills, and lower GAF scores.

Some researchers, such as Horowitz, argue that posttraumatic symptoms may wax and wane with triggers or life events. According to the diagnostic criteria of PTSD in the DSM-IV, the triggers are internal or external cues that symbolize or resemble aspects of the primary stressor. When a person is exposed to triggers, his or her state may change from asymptomatic or subthreshold to symptomatic, or from prodromal mild to severe, and thus meet the diagnostic criteria of D-PTSD. Most researchers and clinicians would probably regard triggers as idiosyncratic, low-grade events with no or only a minimal impact on most people, which, in contrast, lead to major responses in survivors. On the contrary, events leading to reactivation may be regarded as having more impact and being traumatic in a general sense. The notion of reactivation implies that past traumas lower the threshold for certain stressors that may symbolize or resemble the primary stressor, and entail far stronger responses than would be seen in most people. The boundaries between episodes of reactivation as part of D-PTSD and episodes of a new PTSD are rather blurred.

In discussing D-PTSD and concepts such as triggers, reactivating events, and primary stressors, the need for clearer definitions becomes quite apparent. Furthermore, gradients and/or classifications of events may be relevant. More studies are needed to shed light on these issues as well as on the course of the posttraumatic processing. What roles do the primary stressor, the subsequent triggers and aversive events, personality, and the social environment play when full-blown PTSD occurs after a completely asymptomatic period, or after a time dominated by subthreshold levels of PTSD symptoms? If the subthreshold symptoms affect the life of the person to a maladaptive degree yet do not meet the criteria of

PTSD, the person might within certain time limits be diagnosed as having an adjustment disorder.

Many central questions regarding D-PTSD are still unanswered. We do not know much about possible links to aging. More research is needed to expand our current knowledge about the vicissitudes of delayed posttraumatic stress.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Anxiety; Depression and Manic-Depressive Illness.

Further Reading

Arts, P. G. H., Op den Velde, W., Falger, P. R. J., et al. (1996). Late onset of posttraumatic stress disorder in aging resistance veterans in the Netherlands. In: Ruskin, P. E. & Talbott, J. A. (eds.) *Aging and the posttraumatic stress disorder*, pp. 53–76. Washington, D.C.: American Psychiatric Press.

Bryant, R. A. and Harvey, A. G. (2002). Delayed-onset posttraumatic stress disorder: a prospective evaluation. *Australian and New Zealand Journal of Psychiatry* **36**, 2005–2009.

Buckley, T. C., Blanchard, E. B. and Hickling, E. J. (1996). A prospective examination of delayed onset PTSD secondary to motor vehicle accidents. *Journal of Abnormal Psychology* **105**, 617–625.

Eitinger, L. and Strøm, A. (1973). *Mortality and morbidity after excessive stress*. Norway: Universitetsforlaget.

Green, B. L., Lindy, J. D., Grace, M. C., et al. (1990). Buffalo Creek survivors in a second decade: stability of stress symptoms. *American Journal of Orthopsychiatry* **60**, 43–54.

Holen, A. (1990). *A long-term outcome study of survivors from a disaster: the Alexander L. Kielland disaster in perspective*. Oslo, Norway: University of Oslo.

Horowitz, M. J. (1986). *Stress response syndromes*. Northvale: Aronson.

Koren, D., Arnon, I. and Klein, E. (2001). Long term course of chronic posttraumatic stress disorder in traffic accident victims: a three-year prospective follow-up study. *Behavior Research and Therapy* **39**, 1449–1458.

North, C. S., Smith, E. M. and Spitznagel, E. L. (1997). One year follow-up of survivors of a mass shooting. *American Journal of Psychiatry* **154**, 1696–1702.

North, C. S., Pfefferbaum, B., Tivis, L., et al. (2004). The course of posttraumatic stress disorder in a follow-up study of survivors of the Oklahoma City bombing. *Annals of Clinical Psychiatry* **16**, 209–215.

Pomerantz, A. (1991). Delayed onset of PTSD: delayed recognition or latent disorder? *American Journal of Psychiatry* **148**, 1609.

Schnurr, P. P., Lunney, C. A., Sengupta, A. and Waelde, L. C. (2003). A descriptive analysis of PTSD chronicity in Vietnam veterans. *Journal of Trauma and Stress* **16**, 545–553.

Solomon, Z., Kotler, M., Shalev, A. and Lin, R. (1989). Delayed onset PTSD among Israeli veterans of the 1892 Lebanon war. *Psychiatry* **52**, 428–436.

Solomon, Z., Mikulincer, M. and Waysman, M. (1991). Delayed and immediate onset posttraumatic stress disorder. II. The role of battle experiences and personal resources. *Social Psychiatry and Psychiatric Epidemiology* **26**, 8–13.

Stroem, A. (ed.) (1968). *Norwegian concentration camp survivors*. Norway: Universitetsforlaget.

Posttraumatic Stress Disorder, Neurobiology of

J D Bremner
Emory University School of Medicine, Atlanta, GA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J D Bremner, volume 3, pp 186–191, © 2000, Elsevier Inc.

Posttraumatic Stress Disorder: Biology and Phenomenology

Animal Models for Posttraumatic Stress Disorder

Neurotransmitter and Neuropeptide Systems Involved in Stress and Posttraumatic Stress Disorder

Effects of Stress on the Neurobiology of Learning and Memory

Concluding Remarks

Glossary

Amygdala	A small, walnut-shaped brain area in the center of the brain, involved in fear responses.
Cortisol	The stress hormone in humans (glucocorticoid is a more general term, referring to the stress hormone in all animal species).
Extinction	The inhibition of fear responses; felt to be mediated by the medial prefrontal cortex.

Hippocampus	The subcortical brain area involved in learning and memory.
Medial prefrontal cortex	Part of the frontal cortex (brain area just above the eyes) that is involved in emotion and social behaviors; may he involved in posttraumatic stress disorder symptoms.
Magnetic resonance imaging	An imaging technique that can provide detailed maps of brain structure as well as function.
Norepinephrine	A stress neurotransmitter that mediates fight-or-flight responses.
Positron emission tomography (PET)	An imaging technique that uses radiolabeled substances to provide information about brain function.
Posttraumatic stress disorder (PTSD)	A psychiatric disorder, listed in the *Diagnostic and Statistical Manual of Psychiatry*, that follows exposure to an extreme stress that was a threat to life; it is accompanied by intensive thoughts, startle, sleep disturbance, hyperarousal, and other symptoms.

Posttraumatic Stress Disorder: Biology and Phenomenology

PTSD affects 8% of people at some time during their lives. PTSD is characterized by specific symptoms, including intrusive thoughts, hyperarousal, flashbacks, nightmares, sleep disturbance, changes in memory and concentration, and startle responses, that develop following exposure to extreme stress. The symptoms of PTSD are a behavioral manifestation of stress-induced changes in brain structure and function. Stress results in acute and chronic changes in neurochemical systems and specific brain regions, which result in long-term changes in the brain circuits involved in the stress response. Brain regions felt to play an important role in PTSD include the hippocampus, amygdala, and medial prefrontal cortex. Cortisol and norepinephrine are two neurochemical systems that are critical in the stress response. When we refer to the neurobiology of PTSD, we are referring to the aggregate of these stress-induced changes in the brain (Figure 1).

Animal Models for Posttraumatic Stress Disorder

Animal models of stress have been used in the study of stress-related psychiatric disorders, including PTSD. Animals exposed to repeated stress from which they cannot escape develop specific fear-related behaviors that are associated with lasting changes in stress-related neurochemical systems. Behavioral changes related to stress include weight loss and decreases in food intake, active behavior in novel environments,

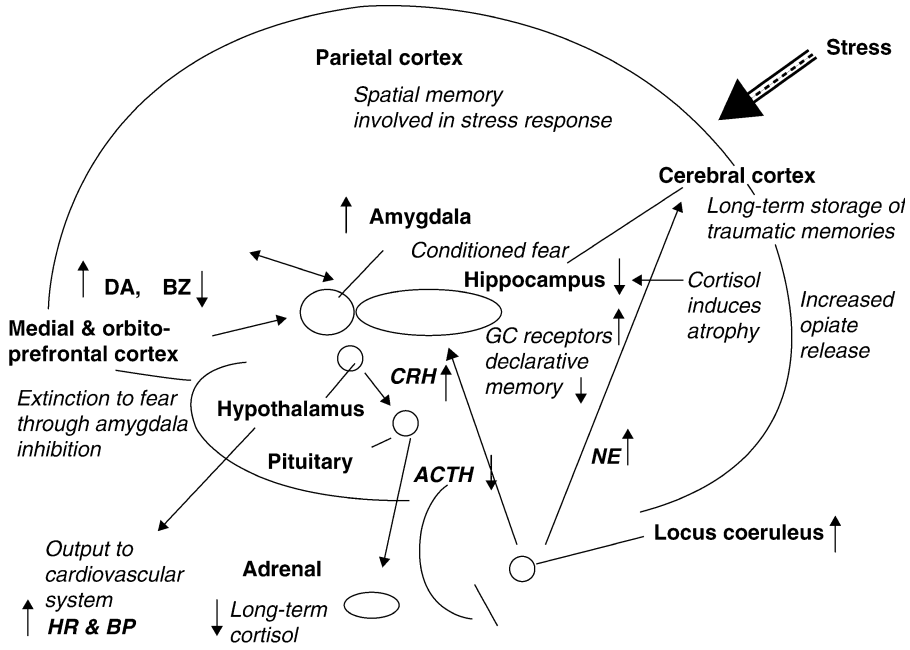


Figure 1 A model for the neurobiology of PTSD. Stress has an effect on multiple brain areas that are responsible for mediating the stress response. These include the amygdala, hippocampus, medial and orbitofrontal cortex, cerebral cortex, hypothalamus, and brain stem. These regions are functionally interrelated and represent important substrates for neurochemical systems involved in the stress response. Long-term changes in the function and structure of these regions lead to symptoms of PTSD. ACTH, adrenocorticotropic hormone; BP, blood pressure; BZ, benzodiazepine; CRH, corticotropin-releasing hormone; DA, dopamine; GC, glucocorticoid; HR, heart rate; NE, norepinephrine.

competitiveness and normal aggressiveness, grooming, play activity, responding for rewards, and self-stimulation of brain reward centers, as well as deficits in memory and attention, sleep disturbances, and increased defecation. These behaviors are similar to stress-related behaviors seen in patients with PTSD.

Neurotransmitter and Neuropeptide Systems Involved in Stress and Posttraumatic Stress Disorder

The locus coeruleus-noradrenergic system plays an important role in the stress response. The majority of norepinephrine cell bodies in the brain are located in a brain-stem site called the locus coeruleus (dorsal pons), with long neurons that project to multiple cortical and subcortical brain sites (e.g., the hippocampus, amygdala, and hypothalamus) for the direct release of the norepinephrine neurotransmitter. A variety of stressors result in an increase in locus coeruleus firing, with increased norepinephrine release in the brain and increased stress-related behaviors (e.g., increased vigilance and fight or flight) that are adaptive in survival. Animals with a history of chronic stress show a potentiated norepinephrine response with reexposure to subsequent stressors. The sensitization of noradrenergic responses may partially underlie the clinical phenomenon of stress sensitization seen in PTSD (e.g., childhood abuse increased the risk of PTSD in veterans who served in Vietnam).

A variety of studies showed long-term dysregulation of the noradrenergic system in PTSD. Kardiner noted in 1941 that veterans with psychiatric conditions related to their war experiences exhibited what appeared to be a hyperresponsiveness of the sympathetic system, manifested by increases in heart rate, blood pressure, sweatiness, irritability, palpitations, vertigo, dizziness, nausea, and syncope. Psychophysiology studies demonstrated an increase in sympathetic responses (heart rate, blood pressure, and galvanic skin response) to traumatic reminders, such as traumatic slides and sounds or traumatic scripts, in PTSD. An increase in norepinephrine and epinephrine was found in 24-h urine samples of patients with combat-related PTSD in some studies but not in others. Plasma levels of norepinephrine were not elevated at baseline in PTSD; however, a potentiated release of norepinephrine was found following exposure to traumatic reminders. Studies of peripheral adrenergic platelets (used as a marker of circulating levels of norepinephrine and epinephrine) have had mixed results. The administration of the adrenergic antagonist yohimbine, which results in an increase in the release of norepinephrine in the brain, resulted in an increase in PTSD-specific symptomatology, as well as a greater release of norepinephrine metabolites in plasma in PTSD patients. Alterations in central metabolic responses to yohimbine were also found in PTSD patients as measured with PET. Yohimbine was associated with a differential effect on brain metabolism in neocortical areas, including the orbitofrontal, temporal, prefrontal, and parietal cortex, with metabolism having a tendency to increase in the controls and decrease in the patients. Given the dose–response effects of norepinephrine on cortical metabolism, these findings are consistent with a potentiated release of norepinephrine in PTSD. In summary, studies are consistent with an increased responsiveness of the norepinephrine system in PTSD.

The mesocortical dopaminergic system, which involves projections from the midbrain to the medial prefrontal cortex, is the most sensitive neural system to mild stressors. Both acute and chronic stressors result in a selective increase in dopamine release and metabolism in the medial prefrontal cortex. The few studies performed in PTSD showed elevations of dopamine and its metabolites in 24-h urine samples. Future work is needed in this important area.

The corticotropin-releasing hormone (CRH)–hypothalamic-pituitary-adrenal (HPA) axis system plays an important role in the stress response. Exposure to stressful situations is associated with a marked increase in cortisol release from the adrenal. Glucocorticoids are important in affecting many of the expressions of the stress response, such as increased gluconeogenesis, inhibition of growth and reproductive systems, and containment of inflammatory responses. Glucocorticoid release from the adrenal is regulated by adrenocorticotropic hormone (ACTH) release from the pituitary, which in turn is regulated primarily by CRH release from the paraventricular nucleus (PVN) of the hypothalamus. In addition to the PVN, CRH is distributed in several brain areas that have been implicated in the behavioral and physiological responses to stress, including the central nucleus of the amygdala, hippocampus, prefrontal and cingulate cortices, locus coeruleus, thalamus, periacqueductal gray, and cerebellum. Intraventricular injection of CRH results in stress-related behaviors (e.g., fearfulness). Stressors early in life may have long-term effects on the CRH–HPA axis, including an increased glucocorticoid response to subsequent stressors. Glucocorticoids have profound effects on neuronal development, which means that early stressors could have lasting effects on development of the brain.

Studies in human subjects have validated the important role of glucocorticoids (cortisol in the human) for the stress response. Stress results in an acute increase in cortisol release, with psychological factors (such as

personal coping, appraisal, and previous stress exposure) modulating cortisol responsiveness. Long-term dysregulation of the HPA axis is associated with PTSD. Baseline levels of urinary cortisol were either decreased or unchanged in chronic PTSD, whereas decreased levels were found in 24-h samples of plasma cortisol levels. Cortisol may be elevated in the more acute phase of PTSD, although further research is needed in this area. A replicated finding has been a supersuppression of the cortisol response after lower doses of dexamethasone suppression test (DST) (0.5 mg), a finding that is the opposite of patients with major depression, who are nonsuppressers with the standard 1-mg DST test. PTSD patients had elevated levels of CRH in the cerebrospinal fluid. Other studies showed an increase in lymphocyte glucocorticoid receptors in PTSD and enhanced cortisol response to stress. One possible explanation of findings to date is an increase in neuronal CRH release, increased central glucocorticoid receptor responsiveness, and resultant low levels of peripheral cortisol due to enhanced negative feedback.

Other neurochemical systems may play a role in symptoms of PTSD. Benzodiazepinelike substances occur naturally in the brain and have been hypothesized to play a role in anxiety and stress. Animal models showed decreased benzodiazepine binding in the frontal cortex with acute and chronic stressors. Consistent with this, neuroimaging studies involving the quantitation of benzodiazepine receptor binding in the human brain showed decreased binding in frontal cortex of patients with PTSD. No difference was found in the anxiety response to the benzodiazepine receptor antagonist flumazenil between patients with combat-related PTSD and controls. Exposure to stress results in an increased release of opiate peptides in the brain associated with the development of an analgesia to pain known as stress-induced analgesia. Opiates cause a decrease in firing from the locus coeruleus; this provides an explanation for the favorable response of hyperarousal symptoms of PTSD to opiates such as heroin. Vietnam veterans with PTSD had a reduced sensitivity to pain during exposure to traumatic reminders in the form of a videotape with combat-related scenes; this was reversible with the opiate antagonist naloxone. Studies showed decreased levels of β-endorphin in the cerebrospinal fluid of patients with PTSD. Stressors result in increased serotonin (5-HT) turnover in the medial prefrontal cortex. Chronic restraint stress results in a decrease in 5-HT1A binding in the hippocampus and in a decrease in 5-HT2 binding in the parietal cortex. Patients with combat-related PTSD had reduced paroxetine-binding sites in platelets (evidence of reduced serotonin reuptake site function). Other studies found no difference in platelet serotonin levels or uptake or in response to probes of 5-HT1A function. Thyroid hormone (thyroxine, T_3) was found to be increased, with a pattern of increased thyrotropin-stimulating hormone response to thyrotropin-releasing hormone in PTSD. Other studies showed increased somatostatin levels in cerebrospinal fluid and decreased levels of neuropeptide Y (NPY) in plasma in PTSD.

Effects of Stress on the Neurobiology of Learning and Memory

There has been considerable interest in alterations in memory function in PTSD. Memory alterations in PTSD take the form of deficits in explicit verbal memory as well as in potentiation of recall of traumatic experiences, dissociative flashbacks, and changes in implicit memory (e.g., fear conditioning). Brain areas, including the hippocampus, amygdala, and medial prefrontal cortex, may mediate these alterations in memory. The hippocampus, a brain area involved in verbal declarative memory, is very sensitive to the effects of stress. Stress in animals was associated with damage to neurons in the CA3 region of the hippocampus with associated deficits in learning and memory. Studies in a variety of animal species suggested that direct glucocorticoid exposure results in decreased dendritic branching and a loss of neurons in the hippocampus. Glucocorticoids appear to exert their effect by increasing the vulnerability of hippocampal neurons to insults such as excitatory amino acids. Other factors, including alterations in serotonin function and decreased levels of brain-derived neurotrophic factor, also probably contribute to stress-related hippocampal damage. Recently, stress has been shown to inhibit neurogenesis, an effect that is reversed by antidepressant treatment.

Studies in PTSD are consistent with abnormalities of the hippocampus. PTSD patients show specific deficits in verbal declarative memory function based on measures that have been validated as sensitive to the loss of neurons in the CA3 region of the hippocampus. Patients with PTSD secondary to combat or childhood abuse also had a smaller volume of the hippocampus based on measurements with magnetic resonance imaging. Deficits in memory were correlated with reduced hippocampal volume in patients with combat-related PTSD. Declarative memory function improved and hippocampal volume increased with antidepressant treatments, including paroxetine and phenytoin. We have hypothesized that stress-induced hippocampal dysfunction may mediate many of the symptoms of PTSD that are related to memory dysregulation, including explicit memory deficits and

fragmentation of memory in abuse survivors. The hippocampus and adjacent cortex plays an important role in binding together information from multiple sensory cortices into a single memory at the time of retrieval. Abnormalities of hippocampal function in PTSD may impair the ability of the hippocampus to recreate a normal memory. Traumatic memories may not be placeable in space and time, leading to fragmentation or amnesia. Traumatic memories may intrude into consciousness without any discernible reason, leading to the symptom of intrusive memories and flashbacks.

The amygdala is involved in memory for the emotional valence of events. The paradigm of conditioned fear has been used as an animal model for stress-induced abnormalities of emotional memory. When an aversive stimulus (e.g., electric shock) is paired repeatedly with a neutral stimulus (bright light), exposure to the conditioned stimulus alone results in an increase in fear responsiveness and an increased startle response. This phenomenon, known as fear conditioning, is mediated by the amygdala. The startle response has been used as a probe of amygdala function in humans. In PTSD patients, the startle response has been shown to be either equal to or greater than that seen in normal subjects. Imaging studies have shown increased amygdala response to external threats, such as fear conditioning or exposure to fearful faces, but not during exposure to internally generated traumatic reminders.

The medial prefrontal cortex has been hypothesized to mediate symptoms of PTSD. The medial prefrontal cortex includes the anterior cingulate gyrus (Brodmann's area 32) and subcallosal gyms (area 25), as well as orbitofrontal cortex, and is functionally and anatomically distinct from the dorsolateral prefrontal cortex. Lesion studies demonstrated that the medial prefrontal cortex is involved in emotion and social behavior. Medial prefrontal cortical areas modulate emotional responsiveness through the inhibition of amygdala function, and dysfunction in these regions may underlie pathological emotional responses in patients with PTSD. Conditioned fear responses are extinguished following repeated exposure to the conditioned stimulus in the absence of the unconditioned (aversive, e.g., electric shock) stimulus. The extinguished memory is rapidly reversible following reexposure to the conditioned–unconditioned stimulus pairing even up to 1 year after the original period of fear conditioning, suggesting that the fear response did not disappear but was merely inhibited. This inhibition appears to be mediated by the medial prefrontal cortical inhibition of amygdala responsiveness. Studies in animals showed that early stress results in decreased brain neurons in the medial prefrontal cortex. Imaging studies in PTSD show a smaller volume of the anterior cingulate. Imaging studies of brain function in PTSD are consistent with dysfunction of the medial prefrontal cortex during the presentation of traumatic cues. Stimulation of PTSD symptoms with the noradrenergic agent yohimbine resulted in a relative failure of activation of metabolism in parts of the medial prefrontal cortex, as well as decreased function in the hippocampus, as measured with the PET assessment of metabolism in PTSD. Exposure to traumatic reminders in the form of traumatic slides and sounds or traumatic scripts was associated with an increase in PTSD symptoms, decreased blood flow in the medial prefrontal cortex and anterior cingulate, and a failure of activation in the anterior cingulate (areas 32 and 24), as measured with PET. Traumatic reminders also resulted in increased activation in the posterior cingulate and motor cortex in two or more studies in PTSD. These areas are involved in visuospatial processing and preparation for action that may be critical to the fight-or-flight response. A decreased function in the hippocampus was also seen in two studies. In summary, dysfunction of the medial prefrontal cortex during exposure to traumatic reminders may underlie the symptoms of PTSD.

Concluding Remarks

Multiple neurotransmitter and neuropeptide systems are involved in the stress response, including CRH, norepinephrine, serotonin, dopamine, endogenous benzodiazepines, and endogenous opiates. These transmitters mutually regulate one another in the execution of the stress response. Chronic stress is associated with long-term alterations in these transmitters and peptides, which translate into changes in neuronal function and structure. Preliminary clinical studies in patients with PTSD are beginning to replicate many of the findings from animal studies. Specific brain regions, including the hippocampus and adjacent cortex, amygdala, and medial prefrontal cortex, are hypothesized to play a role in symptoms of PTSD. Neurotransmitter and neuropeptidal systems involved in stress modulate behavior and memory function through direct actions on these brain regions. Chronic stress, probably through long-term alterations in these systems, can lead to long-term behavioral and memory disturbances. This may represent the mechanism for many of the symptoms of PTSD.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Amygdala; Learning and Memory, Effects of Stress on; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed.

Further Reading

Bremner, J. D. (2002). *Does stress damage the brain? Understanding trauma-related disorders from a mind-body perspective*. New York: W. W. Norton.

Bremner, J. D. (2005). *Brain imaging handbook*. New York: W. W. Norton.

Bremner, J. D., Randall, P. R., Scott, T. M., et al. (1995). MRI-based measurement of hippocampal volume in patients with combat-related posttraumatic stress disorder. *American Journal of Psychiatry* **152**, 973–981.

Bremner, J. D., Southwick, S. M. and Charney, D. S. (1999). The neurobiology of posttraumatic stress disorder: an integration of animal and human research. In: Saigh, P. A. & Bremner, J. D. (eds.) *Posttraumatic stress disorder: a comprehensive text*, pp. 103–143. New York: Allyn & Bacon.

Elzinga, B. M., Schmahl, C. S., Vermetten, E., et al. (2003). Higher cortisol levels following exposure to traumatic reminders in abuse-related PTSD. *Neuropsychopharmacology* **28**(9), 1656–1665.

Rauch, S. L., Whalen, P. J., Shin, L. M., et al. (2000). Exaggerated amygdala response to masked facial stimuli in posttraumatic stress disorder: a functional MRI study. *Biological Psychiatry* **47**(9), 769–776.

Sapolsky, R. M. (1996). Why stress is bad for your brain. *Science* **273**, 749–750.

Vermetten, E. and Bremner, J. D. (2002). Circuits and systems in stress. I: Preclinical studies. *Depression & Anxiety* **15**, 126–147.

Vermetten, E. and Bremner, J. D. (2002). Circuits and systems in stress. II: Applications to neurobiology and treatment of PTSD. *Depression & Anxiety* **16**, 14–38.

Vermetten, E., Vythilingam, M., Southwick, S. M., et al. (2003). Long-term treatment with paroxetine increases verbal declarative memory and hippocampal volume in posttraumatic stress disorder. *Biological Psychiatry* **54**(7), 693–702.

Yehuda, R., Teicher, M. H., Trestman, R. L., et al. (1996). Cortisol regulation in posttraumatic stress disorder and major depression: a chronobiological analysis. *Biological Psychiatry* **40**(2), 79–88.

Posttraumatic Therapy

A M Rasmusson
Veterans Administration National Center for PTSD, West Haven, and Yale University of Medicine, New Haven, CT, USA
C M Monson and P A Resick
Veterans Administration National Center for PTSD and Boston University School of Medicine, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A M Rasmusson and D S Charney, volume 3, pp 192–200, © 2000, Elsevier Inc.

Background: Diagnosis and Symptoms of Posttraumatic Stress Disorder (PTSD) and Acute Stress Disorder (ASD)

Early Interventions

Treatment of PTSD

The Pharmacological Treatment of PTSD

Summary

Glossary

Double-blind study — A research study in which both the research patient and the research clinician are blind or unaware of the type of medication or other intervention being used to treat the patient. This prevents the expectations of the patient and researcher from biasing the research results. In such studies, a research pharmacist or another medical clinician not in direct contact with the patient knows which medication the patient is taking so that appropriate steps can be taken if the patient develops any side effects.

Extinction — The process of learning that a stimulus strongly paired with a traumatic event (e.g., gunshots) is no longer associated with the traumatic event (e.g., combat). This process requires that the person be re-exposed to trauma-related stimuli, and can cause high levels of distress. However, as the person learns that the traumatic event is no longer going to occur in association with that stimulus, anxiety, fear, defensive, and avoidance behaviors subside.

Flashback — Acting or feeling like one is re-experiencing a traumatic event after it is over. A flashback may seem to occur spontaneously or may occur in response to a sight, sound, smell, or other sensation associated with the initial trauma (e.g., in response to hearing a car backfire, seeing a person that resembles the perpetrator of an assault, feeling the ground vibrate, smelling fire). The person experiencing the flashback may retain a partial

awareness of reality or completely lose touch with reality.

Neurotransmitter A chemical released by neurons (nerve cells) in the brain or peripheral nervous system to communicate with other neurons. Neurotransmitters interact with receptors on nearby neurons, or on the neuron from which the neurotransmitters were released, to increase or decrease their activity. Medications used to treat PTSD are known to influence levels or efficacy of the following neurotransmitters or neurotransmitter receptors: serotonin, norepinephrine (noradrenalin), epinephrine (adrenaline), acetylcholine, dopamine, and γ-aminobutyric acid (GABA).

Placebo An inactive pill given to persons participating in research studies to examine the effectiveness of new medications. Some psychiatric symptoms can improve in response to a placebo substance because individuals getting that substance think that the new medication will be helpful. This is called a placebo response. The symptom responses of persons getting a placebo medication are then compared to the responses of persons getting the real medication. Since real medications usually have some undesirable side effects, a real medication must be able to improve symptoms more than a placebo in order for the real medication to be considered effective and worthwhile.

Background: Diagnosis and Symptoms of Posttraumatic Stress Disorder (PTSD) and Acute Stress Disorder (ASD)

Diagnosis of PTSD

Posttraumatic stress disorder (PTSD) is diagnosed when criteria A, B, C, and D set forth in the *Diagnostic and Statistical Manual of Mental Disorders,* fourth edition, text revision (DSM-IV-TR) are met.

To meet criterion A, a person must have experienced or witnessed, or have been confronted with, an event perceived as extremely threatening (criterion A1) in association with an intense feeling of helplessness, horror, or fear (criterion A2). Examples of traumatic events include being physically assaulted, being raped, participating in combat, being a victim of a natural disaster such as an earthquake or tornado, being in a car accident, witnessing a murder, or hearing about the sudden and unexpected homicide of a loved one. Symptoms of PTSD that may develop in response to a traumatic event comprise the B, C, and D criteria. At least one re-experiencing symptom (criterion B), three avoidance symptoms (criterion C), two hyperarousal symptoms (criterion D), and evidence of functional impairment must be present for a diagnosis of PTSD to be made.

Symptoms of PTSD

Re-experiencing symptoms Re-experiencing symptoms of PTSD include bothersome intrusive memories, thoughts, or images of the trauma, nightmares, flashbacks, and emotional or physiological reactions to stimuli associated with the initial trauma.

Avoidance/numbing symptoms Avoidance/numbing symptoms of PTSD include going out of one's way to avoid stimuli associated with the trauma (e.g., not driving after a car accident, avoiding 4th of July fireworks celebrations after experiencing combat trauma, using alcohol or other drugs of abuse to avoid feelings or memories of the trauma), not discussing the traumatic event because of feelings that get stirred up, partial or complete amnesia for the event, disinterest in everyday activities, emotional numbing, and the sense that one's future is dismal or will be cut short.

Hyperarousal symptoms Hyperarousal symptoms of PTSD include hypervigilance, irritability, anger or rage, difficulty concentrating, startle reactions, and sleep disturbances. The latter may include difficulty falling asleep, middle of the night awakening (often in reaction to nightmares), and early morning awakening. It is not uncommon for persons with untreated PTSD to regularly obtain less than 6 h of sleep per night.

Time Course Depending on the duration of PTSD symptoms and the onset of PTSD symptoms relative to the experience of trauma, three PTSD subtypes may be diagnosed. Treatment approaches may vary depending on the timing and course of symptoms. Acute PTSD is diagnosed when symptoms of PTSD remain or develop more than 1 month after the trauma. Symptoms of acute PTSD may last up to 3 months in duration. Chronic PTSD is diagnosed when PTSD symptoms last longer than 3 months, and delayed-onset PTSD is diagnosed when PTSD symptoms develop more than 6 months after the trauma.

Diagnosis of ASD

A diagnosis of acute stress disorder (ASD) requires the presence of at least three symptoms of dissociation (subjective numbing, detachment, or an absence

of emotional responsiveness, an unawareness of surroundings, derealization, depersonalization, or dissociative amnesia), along with one re-experiencing, one hyperarousal, and one avoidance symptom (discussed previously). By definition, these symptoms may last up to 4 weeks from the trauma. The development of ASD predicts the development of PTSD, though not all persons that develop ASD go on to develop PTSD, and not all persons that develop PTSD have experienced a readily diagnosable ASD.

Comorbid Disorders Associated with PTSD

Other psychiatric disorders such as depression, panic disorder, substance abuse, somatization disorder, phobias, and dissociative disorders are commonly associated with PTSD. It is thought that depression and panic disorder when associated with PTSD are neurobiologically distinct from these disorders diagnosed in the absence of trauma exposure. Alcohol, nicotine, or drug abuse may constitute attempts by the patient to maladaptively cope with PTSD symptoms. Thus, the presence of a substance use disorder in a person suffering from PTSD can, in many cases, be construed as an avoidance symptom that has significant secondary adverse effects on the person's adaptation. It is also possible that individuals with comorbid PTSD and substance use disorders have risk factors that predispose them to both. Interestingly, the risk for nicotine dependence is increased by having PTSD, while the risk for the development of PTSD, after trauma exposure, is increased by prior nicotine dependence. Chronic pain and stress-exacerbated medical disorders such as gastric ulcers or irritable bowel syndrome also occur frequently in persons with PTSD. The presence of these comorbid psychiatric and medical disorders can significantly complicate the treatment of PTSD (see later).

Trauma exposure can lead to an intensification of central neurochemical and peripheral neurohormonal responses to perceived threat. Associated state changes facilitate fight or flight reactions at the expense of higher-order cognitive efforts to appraise the actual degree of threat and enact appropriate, considered responses. Patients with PTSD therefore often have problems with impulse control and defensive aggression. Traumatic experiences in childhood and even early adulthood often lead to the development of affect dysregulation, cognitive misappraisals, interpersonal problems, and behaviors that cause self-harm. These symptoms are sometimes considered features of borderline, narcissistic, and antisocial personality disorders, while others consider them symptoms of complex PTSD. Trauma experienced during adulthood may sometimes precipitate the emergence of similar problems along with the development of PTSD even in previously well-functioning individuals. Therapists engaging in the treatment of PTSD therefore should have familiarity with, comfort with, and experience in the management of these defensive, maladaptive patterns of function in order to maintain a therapeutic relationship within which treatment can proceed.

Early Interventions

The risk for developing PTSD after exposure to a traumatic event depends on the type and severity of the trauma, as well as the age, gender, and other premorbid characteristics of the person suffering the trauma. For example, PTSD is thought to occur at higher rates in response to unpredictable or uncontrollable events and in response to interpersonal trauma resulting from the breakdown of expectable psychosocial norms (e.g., rape, genocide, war atrocities, intrafamilial abuse). Research also suggests that a history of childhood trauma may predispose to the development of PTSD in response to adult trauma. Genetic factors may also predispose to the development of PTSD. Finally, the degree of psychosocial support experienced in the aftermath of trauma affects the progression of ASD to PTSD (see later).

Understanding the mechanisms by which the preceding factors contribute to the risk for PTSD will help lead to new methods of PTSD prevention. In the meantime, significant progress has been made in evaluating the efficacy of early interventions aimed at preventing the progression of acute trauma reactions to PTSD. Single-session debriefing therapy was the first such treatment widely employed. During debriefing therapy, trauma victims are educated about typical stress reactions and ways to cope with them. Victims are also encouraged to express feelings and thoughts about the trauma and are informed about other treatments that may be accessed later should the need arise. Controlled trials of debriefing therapy, however, have not only failed to show its efficacy, but also suggest that it might increase the risk for PTSD. Since these studies, research methodologists have emphasized the importance of comparing outcomes of experimental therapies to the natural evolution of trauma symptoms.

Much better outcomes have been found with cognitive-behavioral therapies that employ exposure techniques and/or cognitive restructuring interventions. These interventions have been introduced a few months after trauma exposure and are significantly more effective in preventing PTSD than supportive psychotherapies, the simple repetition of PTSD

symptom assessments, and self-help programs. Importantly, rates of PTSD remittance in response to early cognitive-behavioral interventions are also much higher than natural rates of PTSD remission.

Further research is needed to determine whether even earlier psychological or pharmacological interventions may protect against the development of PTSD, as suggested by preliminary controlled studies of short-term propranolol administration to trauma victims seen initially in an emergency room setting.

Treatment of PTSD

Exposure Therapy

Exposure therapies used to treat PTSD are derived from research into the mechanisms underlying classical and operant conditioning. For instance, it has been fairly well established that traumatic memories are indelibly encoded within the amygdala, a part of the brain that coordinates the hormonal, cardiovascular, behavioral, and cognitive responses to threat in mammals. Within the amygdala, memories of the naturally threatening (i.e., unconditioned) stimulus and the initially nonthreatening stimuli associated with the threatening stimulus become paired at a neuronal level, so that in the future, the previously neutral but now fear-conditioned stimuli alone can elicit fearful defensive responses. Over time, if the individual repeatedly reencounters conditioned stimuli in the absence of the original threatening stimulus, new information about the safety of the conditioned stimuli restrains activation of the amygdala, and the expression of conditioned fear eventually extinguishes. Indeed, it is believed that the success of exposure therapy in PTSD relies in large part on the engagement and enhancement of this natural extinction process.

There are several types of exposure therapy used in the treatment of PTSD. These treatments vary in three ways: (1) the exposure to trauma-related stimuli may be imaginal or involve actual confrontation with feared persons, places, or things; (2) the exposure may be short or long; and (3) the level of emotional arousal stimulated by the exposure may be low or high. In systematic desensitization, patients are instructed to relax, and then are asked to imagine a minimally threatening stimulus for a short period of time. This is repeated several times until the stimulus ceases to cause anxiety. The patient is then asked to imagine a more fearful stimulus for several repetitions until it ceases to elicit anxiety. This process continues until the most threatening situation can be comfortably confronted by the patient. In flooding, the patient starts by confronting the most highly feared stimuli either in the imagination or *in vivo*. The exposure is prolonged and elicits very high levels of anxiety.

Perhaps the most rigorously studied type of exposure therapy is called prolonged exposure (PE). PE is typically conducted in 7 to 10 sessions and involves a combination of imaginal and *in vivo* exposure. After education regarding PTSD, rationale building, and breathing retraining, the therapist and client develop a hierarchy of feared situations for *in vivo* exposure practice between sessions. Then during sessions, which are 90 min in length, imaginal exposure is conducted in which the client is asked to relate the worst traumatic event, in present tense, including as many sensory details as possible. The account is repeated for 45 to 60 min during the session, which is audiotaped. The client listens to the tape between sessions until the next exposure session is taped.

In well-designed studies, PE has been found to result in remission of PTSD in approximately 50% of patients treated. Further studies have suggested that the degree of arousal experienced during exposure therapy influences its outcome. Thus, future research will be needed to identify factors that interfere with achievement of optimal levels of arousal necessary for extinction to occur. Interventions then could be developed to mitigate the impact of such factors. For example, Foa and Meadows have described several anxiety management techniques that could be used by patients with PTSD to manage disabling levels of anxiety induced by exposure to trauma-related cues. Other techniques may help patients who willfully or automatically avoid experiencing necessary levels of arousal and distress during exposure therapy. For patients with neurobiologically mediated tendencies toward over- or underarousal, medications might be used in combination with exposure techniques.

Cognitive Therapy

Cognitive therapy aims to identify, challenge, and restructure automatic systems of meaning that have become attached to traumatic experiences and that interfere with the ability to accurately interpret current events. Cognitive theories of PTSD focus on the effect that various beliefs and schemas (automatic rules) have on emotions and how the memory of the traumatic event may be discrepant with a person's understanding of him- or herself and his or her world view. Cognitive therapy assists the client in examining beliefs about the traumatic event, his or her self, and the world and helps the client challenge possible erroneous and maladaptive assumptions. In several studies, cognitive therapy has been found to be as effective as exposure therapy.

Combination Strategies

Cognitive processing therapy (CPT) Cognitive processing therapy (CPT) utilizes a combination of exposure and cognitive therapy over the course of 12 1-h sessions. CPT was first shown to be efficacious in women with assault-related chronic PTSD, inducing remission of PTSD in about 50% of those who initiated treatment and 80% of those who completed treatment. Since then, CPT has been successfully extended to the treatment of male Vietnam veterans with chronic combat-related PTSD.

During the exposure component of CPT, the patient writes an account of his or her worst traumatic event, reads the account daily, and then reads it to the therapist in session. The patient is encouraged to experience the full range of emotions associated with that event. Most people with PTSD re-experience the traumatic event as only brief sensory images rather than the complete story. Writing the account in detail restores the full event within its greater context. In addition, reviewing the details of the traumatic experience allows the therapist and patient to consider possible assumptions that were made at the time of the trauma and distortions in the patient's retrospective evaluation of the traumatic event that have interfered with recovery. For example, victims of trauma often overestimate the amount of control they had over the event. Self-blame by patients is frequent even when it is obvious to everyone else that the event was not preventable. These cognitive distortions, believed to play a role in the maintenance of PTSD, are referred to as stuck points.

During the subsequent cognitive therapy sessions, patients are taught to systematically challenge their counterproductive thoughts by asking and answering a series of questions and then generating more balanced alternative thoughts. During CPT, new reality-based thoughts begin to supplant the conditioned and overgeneralized automatic beliefs, leading to a marked reduction in re-experiencing, avoidance, and hyperarousal symptoms. Systematically addressing deleterious effects of the traumatic experiences on the victim's self-concept and relationships with others also results in significant reductions in depression and guilt. Indeed, CPT has been found to be modestly more effective than prolonged exposure therapy in reducing depression. It is also more effective in reducing guilt and somatic symptoms comorbid with PTSD.

The therapeutic effects of these evidence-based cognitive-behavioral therapies can be profound and have been shown, in a long-term follow-up study, to persist up to 5 years after treatment.

Eye movement desensitization reprocessing (EMDR) During eye movement desensitization reprocessing (EMDR), a patient tracks the movement of a therapist's finger, while visualizing a traumatic scene and paying attention to any thoughts or feelings provoked by the scene. The therapist takes note of whether the patient's appraisal of the circumstances or meaning of the trauma may be distorted and thereby contribute to negative feelings associated with the incident. This process is repeated until the anxiety generated by the traumatic scene subsides. At this point, the therapist interjects an alternate, and ideally more realistic or constructive, thought about the trauma scene for the patient to consider. This positive association is thought to compete with negative associations and contribute to the extinction of conditioned fear.

It is important to note that recent research has suggested that having the patient follow the movement of the therapist's finger may lessen the intensity of the emotional response to the visualized traumatic material. Thus, having the patient engage in other mildly distracting repetitive movements also may be useful in optimizing the degree of arousal attained during this imaginative exposure exercise and thus may facilitate extinction.

A recent study has found that prolonged exposure therapy results in greater improvements in PTSD symptoms than EMDR, while EMDR did not differ significantly in efficacy from relaxation therapy. A head-to-head comparison between CPT and EMDR has not been conducted.

Treatment of Complicated PTSD

As noted, the presence of comorbid medical and psychiatric conditions can significantly complicate the treatment of PTSD. Although it has been generally believed that substance use disorders need to be addressed before the treatment of PTSD, there have been recent efforts to treat comorbid substance abuse and PTSD simultaneously, with positive effects. States of intoxication as well as withdrawal intensify PTSD symptoms and may compromise the natural biological processes involved in extinction and the development of more reality-based cognitions. In addition, substance dependence results in secondary psychosocial problems (legal, vocational, interpersonal) that further stress and destabilize individuals struggling with PTSD. Comorbid depression also may complicate recovery from PTSD. While both PE and CPT have been shown to significantly decrease depressive symptoms in patients with PTSD, women with PTSD and comorbid major depression have significantly higher self-reported PTSD and depression symptoms at baseline and after treatment with CPT and PE compared to women with PTSD alone. Severe problems with affect regulation frequently

seen in victims of childhood abuse can also present a challenge during PTSD treatment. In addition, chronic pain syndromes associated with PTSD need to be addressed because chronic pain appears to increase PTSD symptoms and vice versa.

Therefore, treatments for PTSD in persons with comorbid psychiatric and medical conditions deserve attention in future research studies. For example, severe dysregulation of affect has been specifically targeted in one therapy format prior to undertaking prolonged exposure therapy in victims of childhood abuse. It has been suggested that comorbid major depression or panic disorder should be addressed pharmacologically prior to undertaking formal treatment of core PTSD symptoms. On the other hand, it is possible that substantial numbers of such patients may be refractory to antidepressants or other medications currently available (see later), or that therapeutic techniques as well as medication administration may helpful. The treatment of chronic pain syndromes associated with PTSD can also be difficult. Standard pain medications such as nonsteroidal anti-inflammatory agents may be ineffective, while tolerance to more effective agents such as opiates can occur; in addition, the possibility that drugs such as opiates may interfere with exposure and cognitive techniques needs to be evaluated. Antidepressants may effectively treat both chronic pain and PTSD in some individuals, but do not benefit all.

The Pharmacological Treatment of PTSD

Over the past several years, significant progress has been made in evaluating the efficacy of psychotropic medications used in the treatment of PTSD. A number of placebo-controlled trials of medication for the treatment of PTSD have been conducted, which has resulted in an increase in the number of agents that can be prescribed with some confidence. In addition, open label studies and case reports continue to suggest other promising medications that likely will be tested in future controlled trials.

Another challenge in the future will be to determine whether psychotropic medications might enhance the efficacy of evidence-based psychotherapies for PTSD. While these psychotherapies are relatively brief and appear to effectively cure PTSD in many individuals, many individuals remain symptomatic after treatment. In addition, otherwise well-controlled trials of these therapies have not controlled for medication use. Thus, we cannot be sure whether medication use might have enhanced or hindered the effects of these psychological treatments in some individuals. Comparisons between the psychological and medication treatments are also difficult because active psychological treatments are generally compared to no treatment, while active medications are compared to nonactive placebo medications that typically lead to some improvement in PTSD symptoms themselves. This could contribute to the greater apparent efficacy of psychotherapies compared to medications.

Antidepressants Used in the Treatment of PTSD

Selective serotonin reuptake inhibitors (SSRIs) Selective serotonin reuptake inhibitors (SSRIs) interfere with the reuptake of serotonin into neurons from which serotonin is released. In addition, SSRIs appear to increase brain levels of neuroactive steroids that reduce anxiety, arousal, and pain. A number of preliminary open label studies, as well as placebo-controlled double-blind studies, have shown that SSRIs such as fluoxetine, sertraline, paroxetine, and citalopram are helpful in treating core symptoms of PTSD. For example, studies of sertraline show approximately 30% reductions in PTSD symptoms. There is some evidence that the efficacy of these agents is greater either in women or in victims of civilian trauma compared to male victims of combat trauma.

Tricyclic antidepressants Tricyclic antidepressants increase brain levels of neurotransmitters such as norepinephrine and serotonin by interfering with their reuptake into neurons after their release. Tricyclic antidepressants may be helpful in some patients with PTSD in the treatment of depression, as well as in the treatment of re-experiencing and hyperarousal symptoms. However, these agents are generally less effective than either monoamine oxidase inhibitors (MAOIs) or SSRIs in the treatment of core symptoms of PTSD. The anticholinergic side effects of tricyclics appear to be less well tolerated in patients with PTSD compared to other patient groups.

Monoamine oxidase (MAO) inhibitors MAO inhibitors are medications that inhibit enzymes that break down monoamine neurotransmitters such as dopamine, serotonin, and norepinephrine. The levels of these neurotransmitters in the brain and peripheral nervous system therefore increase. MAOIs have been shown in double-blind, placebo-controlled studies to be helpful in treating re-experiencing symptoms in PTSD. The use of MAO inhibitors is contraindicated in patients with comorbid substance use because of potential life-threatening interactions with illicit drugs and alcohol. Patients on MAOIs also must adhere to a special diet devoid of foods and beverages with tyramine (an amino acid that interacts with neurotransmitters increased by MAOIs) in order to avoid potentially dangerous side effects. Many

patients are not willing or reliably able to restrict their diet in this manner.

Sleep Medications

A number of medications can be useful in promoting sleep in PTSD. These include the antidepressants trazodone, mirtazapine, and amitriptyline, atypical neuroleptics (see next section), antiadrenergic agents (see later), and drugs that target the γ-aminobutyric acid type A ($GABA_A$) receptor such as the benodiazepines. Problems with tolerance frequently complicate use of the latter, however.

Atypical Neuroleptics Used in the Treatment of PTSD

Neuroleptics act by blocking the effects of neurotransmitters such as dopamine, serotonin, and norepinephrine at their receptors. Traditional neuroleptics primarily block dopamine type 2 (D2) receptors while newer atypical neuroleptics primarily block serotonin type 2 receptors (5-HT2) and sometimes noradrenergic α1 or α2 receptors. Before PTSD was formally recognized as a distinct psychiatric disorder, veterans displaying symptoms of hypervigilance, paranoia, agitation, dissociation, trauma-related auditory hallucinations, and behavioral dyscontrol were sometimes treated with traditional neuroleptics. Neuroleptic use in PTSD was thereafter discouraged due to their poor efficacy, intolerable short-term side effects, and potential for inducing long-term movement disorders. More recently, the newer atypical neuroleptics have been found to be both tolerable and efficacious over the short term in treating severe combat-related PTSD, particularly impulsive aggression, irritability, and sleep disturbance. The long-term effectiveness, tolerability, and long-term potential for side effects associated with these agents have not yet been studied, however.

Other Medications of Potential Use in the Treatment of PTSD

A variety of other medications also may be useful in treating PTSD. However, double blind, placebo-controlled studies of the efficacy of these agents either alone or in combination with other medications have not yet been completed.

Antiadrenergic agents Antiadrenergic medications such as propranolol (which blocks the effects of epinephrine at β-receptors) or clonidine (which inhibits the release of norepinephrine) have been reported to effectively treat hyperarousal and re-experiencing symptoms in PTSD. Large controlled trials of propranolol administered in the immediate aftermath of trauma for the prevention of PTSD are currently under way. Prazosin, a noradrenergic α1 receptor antagonist, has shown some promise in reducing nightmares in PTSD.

Anticonvulsants Preliminary studies suggest that anticonvulsants such as carbamazepine or valproic acid have variable effects on hyperarousal, re-experiencing, and emotional numbing in PTSD. Benzodiazepines (tranquilizers) may be used for the nonspecific symptoms of irritability, anxiety, and insomnia, though tolerance and dependence are often a problem. The use of benzodiazepines in patients with past alcohol or drug abuse problems may be unwise since it may serve to trigger relapse; if prescribed, careful monitoring is necessary. A recent open label trial of topiramate, an antikindling agent that antagonizes certain excitatory glutamate receptors while enhancing the effects of GABA at a novel subtype of the $GABA_A$ receptor, suggests that it may be beneficial and well tolerated in PTSD. Controlled trials of this agent will need to be conducted to confirm its efficacy.

Pharmacological agents under development As our understanding of the underlying neurobiology of PTSD increases, the potential for the development of pharmacological agents that counteract deleterious biological changes associated with PTSD increases. For example, studies have demonstrated an increase in the central nervous system level of the peptide (protein) neurotransmitter called corticotropin-releasing hormone (CRH) in chronically stressed animals and in combat veterans with PTSD. CRH has been shown to confer an increase in stress sensitivity and to increase the expression of defensive behaviors in many animal models. Thus, drugs that interfere with the effects of CRH at its receptors (CRH antagonists) are under development for the possible treatment of PTSD.

Similarly, recent research suggests that blood levels of a substance called neuropeptide Y (NPY) may be decreased in combat veterans with PTSD. Centrally, NPY has been found to counteract the anxiogenic effects of CRH and to be generally anxiolytic. In addition, NPY acts to inhibit the release of norepinephrine in response to stress; therefore, a decrease in blood NPY levels may contribute to the hyperreactivity of the peripheral sympathetic nervous system and the related fight-or-flight behaviors seen in PTSD. Consistent with a role for NPY in modulating hyperreactive stress responses, low plasma NPY levels have been found in combat veterans with PTSD, while higher plasma NPY levels are associated with better performance and decreased dissociation symptoms in military personnel undergoing intense, stressful mock interrogation trials during training. Drugs that target

the NPY system are therefore under development and are being considered for the treatment of PTSD.

Synthetic neuroactive steroids that facilitate action of the inhibitory neurotransmitter GABA at $GABA_A$ receptors are also in development. These steroids have been shown to exert potent anxiolytic, anesthetic, and sedative actions. Recent data suggest that cerebrospinal fluid levels of such neuroactive steroids are low in women with PTSD. Finally, medications that target excitatory, glutamatergic neurotransmitter system receptors such as the *N*-methyl-D-aspartate (NMDA) receptor are being considered for treatment of PTSD. D-cycloserine, for instance, has been shown to enhance the rate of recovery from specific phobias and thus may facilitate extinction of conditioned fear in patients with PTSD.

Summary

Substantial progress has been made in evaluating the efficacy of a variety of psychological and medication treatments for PTSD. Given the demonstrated efficacy of combinations of psychotherapeutic interventions, it is even beginning to appear that chronic PTSD need not be considered chronic. However, there is further work to be done. Many patients with PTSD remain relatively treatment resistant. Biological as well as psychological characteristics of these trauma victims may be important determinants of treatment resistance. Future research needs to better define such factors so that current treatment techniques can be appropriately modified and new treatment approaches developed. Combining pharmacological treatments targeting specific underlying neurobiological abnormalities with psychotherapeutic approaches is likely to be one such new approach. In addition, research aimed at understanding the pathogenesis of PTSD must be continued so that better methods of PTSD prevention and intervention can be developed.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Anxiety; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of.

Further Reading

Albucher, R. C. and Liberzon, I. (2002). Psychopharmacological treatment in PTSD: a critical review. *Journal of Psychiatric Research* **36**, 355–367.

Armony, J. L. and LeDoux, J. E. (1997). How the brain processes emotional information. *Annals of the New York Academy of Science* **281**, 259–270.

Cloitre, M., Koenen, K. C., Cohen, L. R. and Han, H. (2002). Skills training in affective and interpersonal regulation followed by exposure: a phase-based treatment for PTSD related to childhood abuse. *Journal of Consulting and Clinical Psychology* **70**, 1067–1074.

Ehlers, A. and Clark, D. (2003). Early psychological interventions for adult survivors of trauma: a review. *Biological Psychiatry* **53**, 817–826.

Foa, E. B. and Meadows, E. A. (1997). Psychosocial treatments for posttraumatic stress disorder. A critical review. *Annual Reviews of Psychology* **48**, 448–480.

Hien, D. A., Cohen, L. R., Miele, G. M., Litt, L. C. and Capstick, C. (2004). Promising treatments for women with comorbid PTSD and substance use disorders. *American Journal of Psychiatry* **16**, 1426–1432.

Marmar, C. R., Neylan, T. C. and Schoenfeld, F. B. (2002). New directions in the pharmacotherapy of posttraumatic stress disorder. *Psychiatric Quarterly* **73**, 259–270.

Morgan, C. A., III, Rasmusson, A. M., Wang, S., et al. (2002). Neuropeptide Y and subjective distress in humans exposed to acute stress: replication and extension of a previous report. *Biological Psychiatry* **52**, 136–142.

Najavits, L. M., Weiss, R. D., Shaw, S. R. and Muenz, L. R. (1998). "Seeking safety": outcome of a new cognitive-behavioral psychotherapy for women with posttraumatic stress disorder and substance dependence. *Journal of Traumatic Stress* **11**, 437–456.

Nishith, P., Nixon, R. D. and Resick, P. A. (2005). Resolution of trauma-related guilt following treatment of PTSD in female rape victims: a result of cognitive processing therapy targeting comorbid depression? *Journal of Affective Disorders* **86**, 259–265.

Pitman, R. K., Sanders, K. M., Zusman, R. M., et al. (2002). Pilot study of secondary prevention of posttraumatic stress disorder with propranolol. *Biological Psychiatry* **51**, 189–192.

Rasmusson, A., Pinna, G., Weisman, D., et al. (2005). Decreased CSF allopregnanolone levels in women with PTSD correlate negatively with reexperiencing symptoms. Society of Biological Psychiatry, poster presentation. May 19–21, 2005, Atlanta, GA.

Resick, P. A., Nishith, P., Weaver, T. I., Astin, M. C. and Feuer, C. A. (2002). A comparison of cognitive processing therapy with prolonged exposure and a waiting condition for the treatment of chronic posttraumatic stress disorder in female rape victims. *Journal of Consulting and Clinical Psychology* **70**, 867–879.

Ressler, K., J., Rothbaum, B. O., Tannenbaum, L., et al. (2004). Cognitive enhancers as adjuncts to psychotherapy: use of D-cycloserine in phobic individuals to facilitate extinction of fear. *Archives of General Psychiatry* **61**, 1136–1144.

Taylor, S., Thordarson, D. S., Maxfield, L., et al. (2003). Comparative efficacy, speed, and adverse effects of three PTSD treatments: exposure therapy, EMDR, and relaxation training. *Journal of Consulting and Clinical Psychology* **71**, 330–338.

Preeclampsia *See:* Eclampsia and preeclampsia.

Pregnancy – Maternal and Perinatal Stress – Effects of

P Smirnaki and M-A Magiakou
University of Athens School of Medicine, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

Stress, the Stress System and the Hypothalamic-Pituitary-Adrenal Axis in the Nonpregnant State
The Maternal Hypothalamic-Pituitary-Adrenal Axis during Pregnancy
The Hypothalamic-Pituitary-Adrenal Axis of the Fetus
Maternal-Placental-Fetal Stress Physiology
Effects of Prenatal and Perinatal Stress

Glossary

Corticotropin releasing hormone (CRH)	The key hormone of the stress system secreted by the hypothalamus; it stimulates corticotropin production and secretion by the pituitary, which, in turn stimulates cortisol production by the adrenal cortex. During pregnancy, CRH is also produced by the placenta.
Early-life programming	Structural and functional changes in prenatal development caused by early-life environmental factors influence; these persist in the later life or the entire life span of the offspring.
Glucocorticoids	Hormones produced and secreted by the adrenal glands in a baseline circadian fashion and in response to stress; the main endogenous glucocorticoid in humans is cortisol.
Hypothalamic-pituitary-adrenal (HPA) axis	The brain-controlled activation pathway for the secretion of corticosteroid hormones during stress through a hormonal cascade.
Stress	A state caused by the real or perceived challenge to homeostasis, accompanied by various hormonal responses.
Stress system	A complex system of brain circuits and peripheral effectors that are activated during stress. Its main components are the CRH and locus ceruleus–norepinephrine (LC–NE) arousal and autonomic systems and their peripheral effectors, the pituitary-adrenal axis and the limbs of the autonomic nervous system.
Preeclampsia/ eclampsia	Pregnancy associated with disturbances in maternal placental circulation, arterial hypertension, and activation of the acute-phase reaction and blood hypercoagulability. It is a potentially lethal state, frequently associated with spontaneous premature labor and delivery; artificial induction of delivery may be necessary for its control.

Stress, the Stress System and the Hypothalamic-Pituitary-Adrenal Axis in the Nonpregnant State

Living organisms maintain a complex dynamic equilibrium (homeostasis) that is constantly challenged by intrinsic and extrinsic adverse forces or stressors. In humans, the stress response (adaptive responses of the organism to various stressors) results in the activation of a complex repertoire of behavioral and physiological responses. The main components of the stress system are the corticotropin releasing hormone (CRH) and locus ceruleus–norepinephrine (LC–NE) autonomic systems and their peripheral effectors, the pituitary-adrenal axis and the limbs of the autonomic system (Figure 1).

During stress, the hypothalamic-pituitary-adrenal (HPA) axis is activated in response to a variety of stressors, and CRH appears to mediate the metabolic, circulatory, and behavioral changes that are necessary for an effective response. CRH, produced and secreted by the hypothalamus, stimulates adrenocorticotropic hormone (ACTH) production and secretion by the pituitary. In turn, ACTH stimulates the production and secretion of cortisol by the adrenal cortex. The regulation of these hormones is achieved by complex negative feedback mechanisms.

CRH plays an important role in inhibiting gonadotropin-releasing hormone (GnRH) secretion during stress, and via somatostatin it also inhibits growth hormone (GH), thyrotropin-releasing hormone (TRH), and thyrotropin (TSH) secretion, thus suppressing

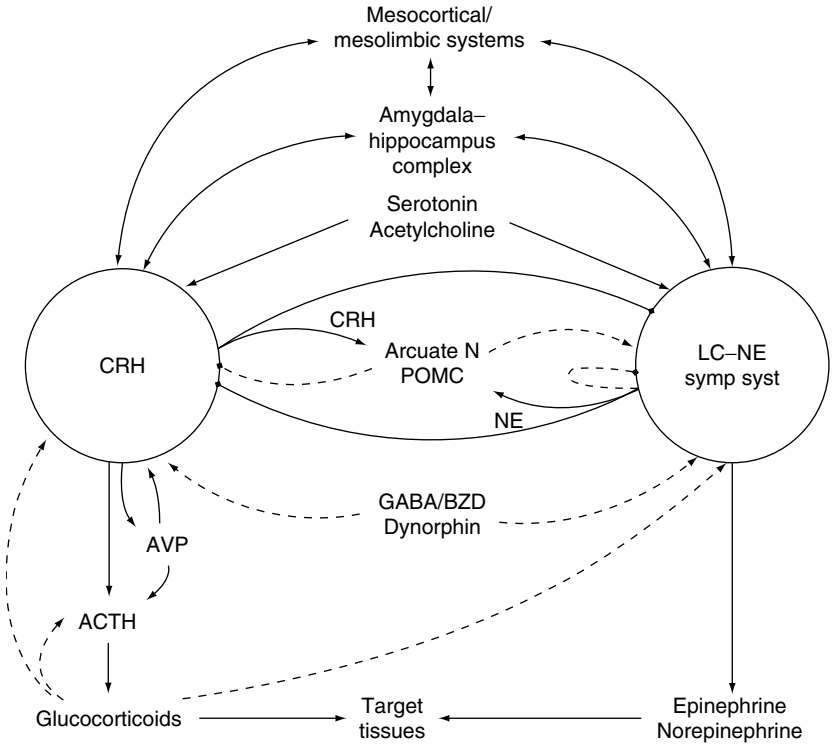


Figure 1 Brain circuits of the stress system. ACTH, adrenocorticotropic hormone; Arcuate N, arcuate nucleus; AVP, arginine vasopressin; BZD, benzodiazepine; CRH, corticotropin releasing hormone; GABA, γ-aminobutyric acid; LC–NE symp syst, locus ceruleus–norepinephrine sympathetic system; NE, norepinephrine; POMC, proopiomelanocortin. From Chrousos, G. P. and Gold, P. W. (1992), The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis, *Journal of the American Medical Association* **267**, 1244–1252.

Table 1 Glucocorticoid functions[a]

Organ system	*Functions*
All fetal organ systems	Maturation (especially of the lung)
CNS	Suppression of pituitary gonadotropins, GH, and TSH
	Inhibition of the CRH, LC–NE, and β-endorphin systems
	Stimulation of the mesocorticolimbic dopaminergic system, the hippocampus, and the CRH perdidergic central nucleus of the amygdala
Adipose tissue	Promotion of visceral adiposity, insulin resistance, dyslipidemia
Cardiovascular system	Enhancement of endothelin-induced vasoconstriction
	Attenuation of endothelium-dependent vasorelaxation
	Affecting renin–angiotensin system receptor density
	Alternation prenatally of the development of cardiac noradrenergic and sympathetic processes
Liver	Regulation of expression of critical hepatic metabolic enzymes, which regulate gluconeogenesis
	Hepatic insulin resistance
Bone	Low-turnover osteoporosis
Autoimmune system	Immunosuppression

[a]CNS, central nervous system; CRH, corticotropin releasing hormone; GH, growth hormone; LC–NE, locus ceruleus–norepinephrine; TSH, thyrotropin.

reproductive, growth, and thyroid functions. Interestingly, these three functions depend on positive catecholaminergic input. On the other hand, the end hormones of the HPA axis, the corticosteroids, play multiple adaptive roles (Table 1). They simultaneously inhibit the CRH, LC–NE and β-endorphin systems and stimulate the mesocorticolimbic dopaminergic system and the CRH-peptidergic central nucleus of the amygdala. In addition, they directly inhibit pituitary gonadotropin, GH, and TSH secretion. They also have direct and insulin-mediated effects on adipose tissue, ultimately promoting visceral adiposity, insulin

resistance, dyslipidemia, and hypertension (metabolic syndrome X) and direct effects on the bone, causing low-turnover osteoporosis. Central CRH via glucocorticoids and catecholamines inhibits the inflammatory reaction, whereas CRH directly secreted by peripheral nerves stimulates local inflammation.

The stress response is meant to be acute or at least of a limited duration. The time-limited nature of this process renders its accompanying antianabolic, catabolic, and immunosuppressive effects temporally beneficial and of no adverse consequences. The chronic, excessive, or deficient activation of the stress system, on the other hand, may lead to abnormal reactivity of the HPA axis, mostly in the form of increased or decreased cortisol secretion. In both cases, pathological adaptation may have detrimental effects on the organism.

The Maternal Hypothalamic-Pituitary-Adrenal Axis during Pregnancy

Pregnancy is a period of hypercortisolism. From weeks 8–10 of gestation on, CRH is also produced by the placenta and secreted primarily into the maternal and also into the fetal circulation. Placental CRH (pCRH) has the same biological activity as hypothalamic CRH (Table 2). pCRH has been postulated to stimulate maternal pituitary ACTH release and cortisol secretion, thereby contributing to pregnancy-associated hypercortisolemia. Paradoxically, cortisol stimulates the synthesis and release of pCRH in a positive feedback regulation fashion. pCRH secretion is also stimulated by inflammatory cytokines, mechanical stretch, and anoxic conditions and therefore by inflammatory and noninflammatory stress, including the stress of preeclampsia or eclampsia. The alterations mentioned result in a shift from normal negative feedback regulation of the maternal HPA axis to a feed-forward mechanism of peripherally produced pCRH, which stimulates adrenal cortisol secretion (Figure 2).

Generally, under normal conditions, circulation pCRH is inactivated to a large extent by a CRH binding protein (CRH-BP). During the second half of pregnancy, plasma CRH levels increase exponentially, whereas the CRH-BP concentration start decreasing at about weeks 33–34 of gestation. The outcome is even higher levels of bioactive CRH, which reaches a peak during labor and delivery and then rapidly decreases postpartum. The increase in CRH levels in the mother's plasma toward the end of pregnancy is normal. HPA axis hormones (CRH; ACTH; corticosteroids; and dehydroepiandrosterone sulfate, DHEAS) are involved in the timing of the onset of the birth process in humans. pCRH causes vasodilation of the fetoplacental circulation by activating nitric oxide synthase and participates in labor by stimulating the secretion of prostaglandin E2a and prostaglandin E2 by the fetal membranes and by promoting the constriction of myometrial cells. At the end of normal pregnancy, the simultaneous stimulation of organ maturation and initiation of parturition constitute a positive effect of a feed-forward mechanism. However, the premature activation of this system may result in preterm labor and delivery. This is supported by evidence of an earlier association between prematurely raised blood pCRH (and decreased CRH-BP) levels in pregnant women who had a preterm delivery. Both preeclampsia and eclampsia are characterized by prematurely elevated levels of pCRH and spontaneously premature labor and delivery.

At the end of parturition, the placental source of CRH is lost; thus, the hypercortisolism of pregnancy is followed by a transient suppression of hypothalamic CRH secretion in the postpartum period. The readjustment of the HPA axis and the postpartum recovery of hypothalamic CRH activity remains poorly characterized and varies among women. The hypothalamic-pituitary-ovarian axis is suppressed

Table 2 Placental corticotropin releasing hormone and its potential physiological functions

Sites of production	*Functions*
Cytotrophoblast, syncytiotrophoblast, amnion, chorion	Maternal hypercortisolism Fetoplacental circulation Timing of labor Fetal adrenal zone stimulation Labor and delivery

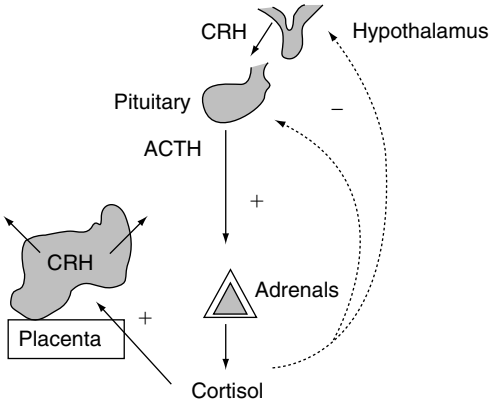


Figure 2 The hypothalmic-pituitary-adrenal axis and placental CRH in human pregnancy. ACTH, adrenocorticotropic hormone; CRH, corticotropin releasing hormone. From Chrousos, G. P., Torpy, D. and Gold P. W. (1998), Interactions between the hypothalamic-pituitary-adrenal axis and the female reproductive system: clinical implications, *Annals of Internal Medicine* **129**, 229–240.

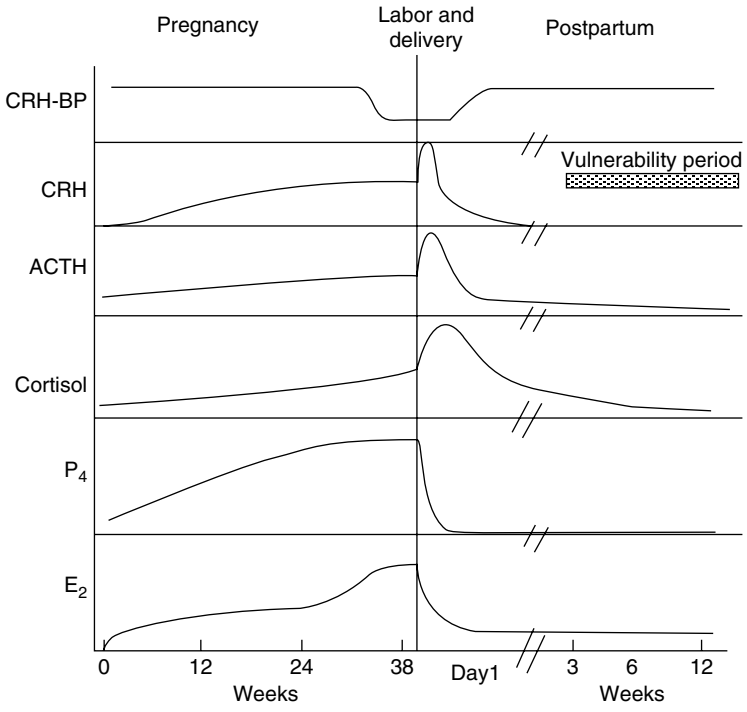


Figure 3 Hormonal changes and period of increased vulnerability to mood disorders and autoimmune phenomena during pregnancy and postpartum. ACTH, adrenocorticotropic hormone; CRH, corticotropin releasing hormone; CRH-BP, CRH binding protein; E_2, 17β-estradiol; P_4, progesterone. From Chrousos, G. P., Torpy, D. and Gold P. W. (1998), Interactions between the hypothalamic-pituitary-adrenal axis and the female reproductive system: clinical implications, *Annals of Internal Medicine* **129**, 229–240.

postpartum and does not recover for approximately 4–6 weeks. Both hypothalamic CRH hyposecretion and abrupt estrogen withdrawal may explain the postpartum blues or depression and autoimmune phenomena frequently seen during this period (Figure 3). Estrogen administration has been shown to ameliorate postpartum depression, probably by returning to normal the secretion of the suppressed CRH neuron.

During pregnancy, ACTH and cortisol continue to be secreted in a pulsatile and circadian fashion, whereas the secretion of pCRH is pulsatile but not circadian. The lack of circadian correlation between CRH and ACTH/cortisol is not surprising because the major part of CRH comes from the placenta, is bound to CRH-BP, and has limited action on the maternal corticotrophs. Moreover, another secretagogue, possibly arginine vasopressin (AVP), released at the median eminence from terminals of the parvocellular neurons, may be important in regulating pulsatile and circadian ACTH secretion during this period.

The Hypothalamic-Pituitary-Adrenal Axis of the Fetus

Fetal cortisol is important for the maturation of virtually all fetal organ systems. The importance of the increase in fetal cortisol secretion for the maturation of the fetal lung and other systems is well established. The fetal HPA axis is regulated through negative feedback starting early in gestation. In a later stage of pregnancy, pCRH enters the fetal circulation via the umbilical vein. Maternal HPA axis hormones also stimulate the placenta to produce CRH. pCRH stimulates the fetal HPA axis to produce and secrete ACTH, cortisol, and androgens (DHEAS, which at the placenta is converted to estriol). Subsequently, fetal cortisol enters the placental circulation via the umbilical vessels and further stimulates the production of pCRH. This way, the fetal HPA axis is also regulated by a feed-forward mechanism at the end of pregnancy. This results in a large increase in cortisol through which maturation of the fetal organs is enhanced.

Maternal-Placental-Fetal Stress Physiology

It is most likely that maternal stress influences the fetus though (stress) hormones. The following three mechanisms have been proposed, which may operate simultaneously and may amplify one another's effects.

1. Reduced uteroplacental blood flow. Corticosteroids and catecholamines exert strong effects on peripheral blood vessels, and the placenta is replete

with receptors for these hormones. The activation of the sympathetic nervous system by stress may lead to reduced blood flow to the uterus and fetus, and it may contribute to fetal growth restriction.

2. Transplacental transport of maternal stress hormones. The human fetus is relatively protected against direct exposure to high cortisol concentrations. In the placenta, 50–90% of maternal cortisol is converted to the biologically inactive cortisone by the enzyme 11β-hydroxysteroid dehydrogenase (11β-HSD2). This enzyme protects the fetus for most of the pregnancy. However, under specific conditions (e.g., if the maternal cortisol concentration is very high, if the activity of 11β-HSD2 is reduced or impaired due to interindividual differences, or when the placenta is immature in early pregnancy or poorly functioning), maternal cortisol is not completely inactivated in the placenta and the fetus is directly exposed to its detrimental actions. Interestingly, in normal pregnancy, a few days before normal labor the 11β-HSD2 spontaneously decreases to further prepare the fetus for the extra-uterine life.

3. Secretion of CRH to the fetus. A diminished supply of nutrients and oxygen (hypoxemia) for any reason, including infection, inflammation, preeclampsia, eclampsia, starvation, or placental insufficiency, may cause a stress response in the fetus. This involves increased secretion of pCRH that, in turn, contributes to the feed-forward mechanisms at either side of the placenta and accelerates labor and delivery.

Effects of Prenatal and Perinatal Stress

The stress-induced activation of the HPA axis during pregnancy can be expected to have serious negative implications for the mother (pregnancy complications, preterm delivery, problems in lactation, and increased vulnerability to mood disorders and autoimmune phenomena postpartum), the fetus's development (intrauterine growth retardation, immaturity, neurological damage, etc.), and subsequent sequelae in the adult life of the newborn infant.

Maternal Stress: Pregnancy Complications and Preterm Delivery

Recent well-controlled human studies indicated that pregnant women with high stress and anxiety levels are at increased risk for spontaneous abortion, preeclampsia, and preterm labor. The HPA axis hormones (CRH, ACTH, corticosteroids, and DHEAS) are involved in the timing of the onset of birth in humans (rapid increase in free pCRH during the last 2–4 weeks of pregnancy). pCRH initiates, through an increase in DHEAS (the precursor of estrogens), a cascade of events that may lead to increased uterine activity and eventually labor and delivery. The fetal DHEAS reaches the placenta, where it is converted to estrone (E_1) and estradiol (E_2) and, after 16-hydroxylation in the fetus's liver, also to estriol (E_3). The increase in E_2 and E_3 contributes to the onset of term and preterm delivery (through the stimulation of oxytocin, oxytocin receptor, gap junction, and prostaglandin synthesis).

Maternal stress and fetal stress or hypoxia increase maternal and/or fetal plasma cortisol levels and, because of its positive stimulatory effect, the latter causes an increase in pCRH. Maternal and fetal stresses are also associated with an increase in NE, angiotensin II, and AVP, all compounds that stimulate CRH release in the hypothalamus as well as in the placenta. Fetal distress and preterm delivery may also be driven by amniochorionic and maternal systemic inflammation. Infections and high levels of interleukins may also stimulate CRH secretion and the stress system. Women with preterm delivery, especially those with infections, show an increase in cytokines such as interleukin (IL-)1, tumor necrosis factor (TNF-)α, and IL-6, which is also secreted by the placenta in response to IL-1 and TNF-α. The activation of the cytokine network may lead to increased placental apoptosis and preterm delivery.

Under abnormal circumstances (preeclampsia, eclampsia, threatened preterm delivery, or maternal stress), all these alterations may be initiated prematurely. In mothers who deliver preterm, the hormone increases might be noticeable even in the second trimester, indicating that the trajectory toward preterm delivery may start early in pregnancy and may be recognizable, potentially leading to the development of preventive interventions. This issue is still in need of further investigation. Potent CRH antagonists might be used as labor retardants.

Maternal Stress and Lactation

Lactating women interact more positively with their babies, and nursing mothers (vs. bottle-feeding mothers) are more likely to describe positive mood states and less anxiety. Successful lactation involves the oxytocin-mediated milk ejection reflex and neuroendocrine adaptations of the HPA and gonadal axes. A recent study conducted in breast-feeding women reported decreased plasma ACTH, cortisol, AVP, epinephrine, and glucose responses to treadmill stress tests, suggesting that, under normal conditions, during lactation there is a selective inhibition of the hypothalamic stress response (stress hyporesponsive period). Central actions of oxytocin have been implicated as a cause of the downregulation of HPA

axis reactivity. On the other hand, it has long been known that stress and hormonal products of the HPA axis can influence the release and actions of oxytocin. Milk ejection is inhibited by stressful experiences, suggesting that oxytocin secretion in women could also be inhibited by stress.

Maternal Stress and Increased Vulnerability to Mood Disorders and Autoimmune Phenomena Postpartum

Epidemiological studies indicate that childbearing represents a period of vulnerability to depressive symptomatology in women. Postnatal depression is a significant and serious illness with a reported incidence of approximately 14%, and a further 30% of women experience an adjustment disorder and anxiety following childbirth. The impact of postpartum depression and psychosis on family life is pervasive, and its capacity for causing long-term harm in children is disturbing. The *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.; DSM IV) indicates that postpartum disorders are distinguished not by their phenomena but by their timing. Any major depressive, manic, or mixed episode; bipolar I disorder or bipolar II disorder; or brief psychotic disorder should get a postpartum specifier if it occurs within the postpartum period.

The abrupt CRH withdrawal once the placenta is delivered and its impact on the mother's HPA axis equilibrium may be a major contributing factor to the development of postpartum blues or depression. It is possible that the development of postpartum blues or depression is related to the prolonged or exaggerated suppression of hypothalamic CRH during the postpartum period because the occurrence of postpartum depression has been found to be related to fluctuations in the levels of glucocorticoids. The resetting of the HPA axis and postpartum recovery of the CRH activity remains poorly characterized and varies among women. Moreover, stress affects glucocorticoid levels and can influence estradiol levels in nonpregnant and pregnant females. The sudden withdrawal of estradiol following parturition is also considered an important pathophysiological mechanism of postpartum disorders. In support of this hypothesis, estrogen is an effective treatment for postnatal depression. The newborns of mothers with prepartum and postpartum depressive symptoms have elevated cortisol and norepinephrine levels and lower dopamine levels.

The transient suppression of hypothalamic CRH secretion in the postpartum period might also explain the increased vulnerability to autoimmune phenomena at this period. A defective HPA axis might lead to relative resistance to infections and neoplastic disease and increased susceptibility to autoimmune and inflammatory disease, such as Hashimoto's thyroiditis, rheumatoid arthritis, and multiple sclerosis.

Prenatal Stress and Its Effects on Fetal Development

Fetal cortisol levels are increased in human fetuses showing intrauterine growth retardation or in pregnancies complicated by preeclampsia, implicating endogenous cortisol in retarded fetal growth. Chronic maternal stress is associated with reduced fetal heart rate variability, which is an indicator of fetal maturation. Most studies assessing the impact of both acute and chronic stress in the fetus have measured changes in heart rate and fetal movements. Experimental studies have demonstrated a relation between specific episodes of maternal psychological stress and exacerbation of fetal asphyxia *in utero*. Experiencing stress during pregnancy leads to the birth of babies with low birth weight, congenital malformations, or growth retardation (reduced head circumference, in particular).

Stress also increases the risk of the infant's developing diseases in subsequent periods of life. Antenatal glucocorticoid administration has been linked with higher blood pressure in adolescence and subtle effects on neurological function, including reduced visual closure and visual memory. Multiple doses of antenatal glucocorticoids given to women at risk of preterm delivery were associated with reduced head circumference and an increased risk of externalizing behavior problems, distractibility and inattention in the offspring.

The Long-Term Programming Hypothesis

The importance of a good quality of psychoaffective communication between mother and child during pregnancy has been shown to be decisive for fetal growth, the perinatal period, and further development of the child. Early-life environmental factors influence prenatal development and may cause structural and functional changes, which persist for the life span. This organizational phenomenon is termed early-life programming or fetal programming of adult disease. This hypothesis proposes that a stimulus or insult acting during critical periods of growth and development may permanently alter tissue structure and function. Genes and environment are no longer considered to exert separate influences. Development is viewed not as a gradual elaboration of an architectural plan preconfigured in the genes but, rather, as a dynamic interdependency of genes and

environment, characterized by a continuous process of interactions in a space- and time-dependent manner. Evidence from both human and animal studies suggests that manipulating the environmental experience of the fetus can induce many diseases of adult life. Programming agents seem to include nutrients, growth factors, cytokines, and hormones, all of which can be altered by stress.

Numerous studies, initially in the United Kingdom and then worldwide, revealed an association between lower birth weight and the subsequent development of the common cardiovascular and metabolic disorders of adult life, notably hypertension, insulin resistance, type 2 diabetes, hyperlipidemia, and atherosclerotic cardiovascular diseases. Thus, early-life events altering birth weight are important predictors of adult morbidity. In addition, postnatal catch-up growth also appears to be predictive of the risk of adult cardiovascular disease, suggesting that it is the restriction of intrauterine growth (caused by overexposure to glucocorticoids) that is pathophysiologically important.

The human HPA axis also appears to be programmed by the early-life environment. Higher plasma and urinary glucocorticoid levels are found in children and adults who were of low birth weight. Two major environmental hypotheses have been proposed to explain the mechanism by which low birth weight is associated with adult disease: fetal nutrition and overexposure of the fetus to glucocorticoids. There is no doubt that in developing countries maternal diet can affect birth weight and may be extremely important in mediating intergenerational effects on birth weight and adult size. Glucocorticoid administration during pregnancy is well documented to both reduce offspring birth weight and alter the maturation of organs (hence, their use to accelerate fetal lung maturation in premature labor). Moreover, antenatal exposure to endogenous or exogenous glucocorticoids produces permanent hypertension, hyperglycemia, hyperinsulinemia, altered behavior, and neuroendocrine responses throughout the life span (in animal experiments).

Fetal undernutrition and fetal glucocorticoid overexposure may be linked by a common mechanism. Processes underlying fetal programming include the determination of the set point of the HPA axis and of tissue glucocorticoid receptor (GR) expression. The molecular mechanisms may reflect permanent changes in the expression of specific transcription factors, key among which is the GR itself. The differential programming of the GR in different tissues reflects effects on one or more of the multiple tissue-specific alternate first exons/promoters of the GR gene. The proposed mechanisms by which overexposure to glucocorticoids prenatally lead to cardiovascular and metabolic disorders of adult life are as follows:

- Hypertension and cardiovascular disease: Prenatal glucocorticoid exposure leads to irreversible reductions in nephron number, fetal and adult vascular responses to vasoconstrictors, enhancing endothelin-induced vasoconstriction and attenuating endothelium-dependent vasorelaxation, indicating microvascular dysfunction. Renin–angiotensin system receptor density and tissue synthesis are also affected. Prenatal glucocorticoid exposure also alters the development of cardiac noradrenergic and sympathetic processes.
- Liver and pancreas: Prenatal overexposure to exogenous or endogenous glucocorticoids programs permanent hyperglycemia – particularly hyperinsulinemia – in the adult offspring, effects that are confined to the last third of gestation. Glucocorticoids regulate the expression of critical hepatic metabolic enzymes, notably phosphoenolpyruvate carboxykinase (PEPCK), which catalyses a rate-limiting step in gluconeogenesis. Prenatal undernutrition impairs pancreatic β-cell development, reducing β-cell mass and causing glucose intolerance. Fetal pancreatic insulin content correlates inversely with fetal corticosterone levels.
- Adipose tissue: Antenatal corticosteroid exposure seems to program fat metabolism, causing a marked increase in GR expression selectively in visceral adipose tissue, which contributes to both adipose and hepatic insulin resistance. Leptin concentrations in human fetal cord blood correlate directly with body weight and adiposity at birth.

Altered Regulation of Hypothalamic-Pituitary-Adrenal Axis after Prenatal Stress and Its Influence on the Fetal Central Nervous System

There is abundant evidence of an overactive and dysregulated HPA axis in the offspring of prenatally stressed animals. Prenatally stressed animals have higher basal blood glucocorticoid levels, and when exposed to a stressful stimulus, their hormonal response is more rapid, stronger, and more prolonged, indicating that the normal negative feedback mechanism of the HPA axis is heavily impaired. In addition, alterations have been found in almost all known regulatory and neurotransmitter systems in the brain, including the opioid, cholinergic, serotonergic, dopaminergic, GABA-ergic, and noradrenergic systems. These alternations could underlie some of the behavioral and physiological effects observed after prenatal stress and may further affect later stress responses.

Three mechanisms can explain the overactivity and altered feedback regulation of the HPA axis (increased levels of HPA axis hormones affect fetal central nervous system development and, especially, brain GR development): (1) the downregulation of receptors in the hippocampus, hypothalamus, pituitary, or adrenal glands; (2) the decreased sensitivity of the receptors at any of these levels; and (3) the alteration in CRH levels or the altered levels or affinity of plasma-binding proteins. The maternal HPA axis is involved in early neuroendocrine programming of the offspring.

The hippocampus appears to be particularly vulnerable to insult during early development because its pyramidal neurons contain a high concentration of GRs that are highly sensitive to either hypercortisolemia caused by severe stress or to exposure to exogenous glucocorticoids. Another neural structure that may contribute to increased responsiveness of the HPA axis following prenatal stress is the amygdala, a key locus for its effects on fear and anxiety. The content of CRH appeared to be substantially increased in the amygdala of prenatally stressed rats compared to levels in control animals. These increased levels of CRH are associated with increased emotionality, fear, and anxiety-like behaviors and facilitate stress-induced behavior and autonomic activation.

Moreover, prenatal glucocorticoid exposure contributes to schizoaffective, attention deficit hyperactivity, and extrapyramidal disorders. Stressful events in the second trimester of human pregnancy have been associated with increased incidence of offspring schizophrenia.

See Also the Following Articles

11β-Hydroxysteroid Dehydrogenases; Corticotropin Releasing Factor (CRF); Hypertension; Metabolic Syndrome.

Further Reading

Austin, M. P., Leader, L. R. and Reilly, N. (2005). Prenatal stress, the hypothalamic-pituitary-adrenal axis and fetal and infant neurobehaviour. *Early Human Development* **81**, 917–926.

Buitelaar, J. K., Huizink, A. C., Mulder, E. J., et al. (2003). Prenatal stress and cognitive development and temperament in infants. *Neurobiology of Aging* **24**, S53–S60.

Carter, C. S., Altemus, M. and Chrousos, G. P. (2001). Neuroendocrine and emotional changes in the postpartum period. *Progress in Brain Research* **133**, 241–249.

Chrousos, G. P. (1999). Reproductive placental corticotropin-releasing hormone and its clinical implications. *American Journal of Obstetrics and Gynecology* **180**, S249–S250.

Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. *Journal of the American Medical Association* **267**, 1244–1252.

Chrousos, G. P., Torpy, D. and Gold, P. W. (1998). Interactions between the hypothalamic-pituitary-adrenal axis and the female reproductive system: clinical implications. *Annals of Internal Medicine* **129**, 229–240.

de Weerth, C. and Buitelaar, J. K. (2005). Physiological stress reactivity in human pregnancy – a review. *Neuroscience and Biobehavioral Reviews* **29**, 295–312.

Drake, A. J. and Walker, B. R. (2004). The intergenerational effects of fetal programming: non-genomic mechanisms for the inheritance of low birth weight and the cardiovascular risk. *Journal of Endocrinology* **180**, 1–16.

Fowden, A. L., Giussani, D. A. and Forhead, A. J. (2005). Endocrine and metabolic programming during intrauterine development. *Early Human Development* **81**, 723–734.

Halbreich, U. (2005). The association between pregnancy processes, preterm delivery, low birth weight, and postpartum depressions – the need for interdisciplinary integration. *American Journal of Obstetrics and Gynecology* **193**, 1312–1322.

Kaiser, S. and Sachser, N. (2005). The effects of prenatal social stress on behaviour: mechanisms and function. *Neuroscience and Biobehavioral Reviews* **29**, 283–294.

Magiakou, M. A. and Chrousos, G. P. (2005). Biological basis of stress-related diseases. In: Antoniou, A. S. G. & Cooper, C. L. (eds.) *Research companion to organizational health psychology*, pp. 70–86. Cheltenham, UK: Edward Elgar.

Magiakou, M. A., Mastorakos, G., Rabin, D., et al. (1996). Hypothalamic corticotropin releasing hormone suppression during the postpartum period: implications for the increase of psychiatric manifestations during this time. *Journal of Clinical Endocrinology and Metabolism* **81**, 1912–1917.

Magiakou, M. A., Mastorakos, G., Rabin, D., et al. (1996). The maternal hypothalamic pituitary-adrenal axis in third trimester human pregnancy. *Clinical Endocrinology* **44**, 419–428.

Majzoub, J. A. and Karalis, K. P. (1999). Placental corticotropin-releasing hormone: function and regulation. *American Journal of Obstetrics and Gynecology* **180**, S242–S246.

Mulder, E. J. H., Robles de Medina, P. G., Huizink, A. C., et al. (2002). Prenatal maternal stress: effects on pregnancy and the (unborn) child. *Early Human Development* **70**, 3–14.

O'Regan, D., Welberg, L. L. A. M., Holmes, M. C., et al. (2001). Glucocorticoid programming of pituitary-adrenal function: mechanisms and physiological consequences. *Seminars in Neonatology* **6**, 319–329.

Seckl, J. R. (2001). Glucocorticoid programming of the fetus; adult phenotypes and molecular mechanisms. *Molecular and Cellular Endocrinology* **185**, 61–71.

Seckl, J. R. (2004). Prenatal glucocorticoids and long-term programming. *European Journal of Endocrinology* **151**, U49–U62.

Van den Bergh, B. R. H., Mulder, E. J. H., Mennes, M., et al. (2005). Antenatal maternal anxiety and stress and the neurobehavioural development of the fetus and child: links and possible mechanisms: a review. *Neuroscience and Biobehavioural Reviews* **29**, 237–258.

Weinstock, M. (2005). The potential influence of maternal stress hormones on development and mental health of the offspring. *Brain, Behavior and Immunity* **19**, 296–308.

Welberg, L. A. M. and Seckl, J. R. (2001). Prenatal stress, glucocorticoids and the programming of the brain. *Journal of Neuroendocrinology* **13**, 113–128.

Premenstrual Dysphoric Disorder

S Nowakowski
San Diego State University/University of California, San Diego, CA, USA
P Haynes
University of Arizona, Tucson, AZ, USA
B L Parry
University of California, San Diego, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Haynes and B L Parry, volume 3, pp 201–207, © 2000, Elsevier Inc.

Diagnostic Issues
Etiology

Glossary

Circadian oscillator — A pacemaker located in the superchiasmatic nucleus of the hypothalamus with an output that repeats about once per day.

Dysphoric mood — A symptom of depression; feeling sad, despondent, discouraged, unhappy, anxious, or irritable.

Entrainment — The process by which the circadian oscillator becomes synchronized with the external environment.

Follicular phase — Also called the proliferative phase or estrogenic phase, the follicular phase of the menstrual cycle lasts from the first day of menstruation until ovulation. As assessed by the secretion of luteinizing hormone (LH), gonadotropin-releasing hormone (GnRH) is secreted from the hypothalamus in a pulsatile manner at a frequency of once per 90 min during the early follicular phase and increasing to once per 60–70 min. Follicle-stimulating hormone (FSH) is secreted by the anterior pituitary, and its concentrations in plasma increase from the first day of the cycle. The rise in FSH levels can be attributed to a decrease in progesterone and estrogen levels at the end of the previous cycle and the subsequent removal of inhibition of FSH by these ovarian hormones. FSH stimulates the development of 15–20 follicles each month and stimulates ovarian follicular secretion of estradiol. Estrogen levels peak toward the end of the follicular phase of the menstrual cycle, at which point estrogen exerts positive feedback on the hypothalamic-pituitary-gonadotrope system, thereby triggering the preovulatory LH surge.

Luteal phase — Approximately the second half of the menstrual cycle occurring after ovulation during which there is a decrease in frequency and increase in amplitude of time, ovarian steroid, and pulsatile release LH (and presumably GnRH) secretion. The luteal cells of the corpus luteum produce progesterone, the plasma levels of which peak by the middle of the luteal phase (days 20–22). By day 28, the corpus luteum has regressed with a consequent decrease in the output of estrogen and progesterone. The latter causes local ischemia and subsequent necrosis of the endometrium (i.e., menstrual flow).

Zeitgeber — German for time giver; a regular, external cue that synchronizes the internal clock to environmental cues; light is the most powerful zeitgebers in humans.

Premenstrual dysphoric disorder (PMDD) is a clearly defined syndrome that is often referred to as premenstrual syndrome (PMS) among the nonmedical population. As delineated in the *Diagnostic and Statistical Manual*, 4th edition (DSM-IV-TR), five symptoms are required to make the diagnosis of

Table 1 Diagnostic criteria for premenstrual dysphoric disorder

A. In most menstrual cycles during the past year, at least five of the following symptoms must have been present most of the time during the last week of the luteal phase and disappear within a week after the onset of menstruation. At least one of the symptoms must be one of the first four.

1. Feeling sad, hopeless, or self-deprecating
2. Feeling tense, anxious or on edge
3. Marked lability of mood interspersed with frequent tearfulness or increased sensitivity to rejection
4. Persistent irritability, anger, and increased interpersonal conflicts
5. Decreased interest in usual activities
6. Difficulty concentrating
7. Feeling fatigued, lethargic, or lacking in energy
8. Marked changes in appetite, which may be associated with binge eating or craving certain foods
9. Hypersomnia or insomnia
10. A subjective feeling of being overwhelmed or out of control
11. Physical symptoms such as breast tenderness or swelling, headaches, or sensations of bloating or weight gain

B. The symptoms must cause an obvious and marked impairment in the ability to function socially or occupationally.

C. The symptoms may be superimposed on another disorder but are not merely an exacerbation of the symptoms of another disorder, such as major depression, panic disorder, dysthymia, or a personality disorder.

D. Criteria A, B, C, and D must be confirmed by prospective daily self-ratings during at least two consecutive symptomatic cycles. (The diagnosis may be made provisionally prior to this confirmation.)

From the American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders,* 4th edition, text revision. Washington, D.C.: American Psychiatric Association, with permission.

PMDD, and at least one symptom must be a moderate to severe mood symptom (see Table 1). The criteria for PMDD are more stringent and delineate increased specificity and severity of premenstrual symptoms with an emphasis on affective symptoms (see Table 1). The criteria for PMDD are stringent and delineate increased specificity and severity of premenstrual symptoms with an emphasis on affective symptoms (see Table 1). Symptoms typically begin in the late luteal (premenstrual) phase of the menstrual cycle, which is associated with declining estrogen and progesterone levels from the degenerating corpus luteum. Premenstrual symptoms typically remit within the first few days of the follicular phase, when estrogen levels increase as a result of follicular growth. The diagnosis of PMDD requires documentation of luteal phase timing of symptoms and a symptom-free interval from approximately day 4 of menses to the onset of ovulation. The National Institute of Mental Health Premenstrual Syndrome Workshop (1983) specified that a 30% change in symptom severity during the luteal phase is required for a PMDD diagnosis. In addition, a diagnostic requirement is that symptoms must cause functional impairment.

Diagnostic Issues

R. T. Frank first identified physical, psychological, and behavioral changes corresponding to monthly changes in reproductive hormones in 1931. However, public awareness of the concurrent change in mood with phases of the menstrual cycle has increased dramatically over the past two decades. The media has disseminated a wide body of information shaping a societal notion of what is popularly termed premenstrual syndrome (PMS). Definitions for PMS have varied widely. Typically, PMS refers to a large number of symptoms, including irritability, tension, fatigue, dysphoria, distractibility, impaired motor coordination, changes in eating and sleeping, and libido changes, which occur in the late luteal phase and remit after the beginning of menstruation. Up to 90% of women experience some premenstrual symptoms, and 50% of women may experience many symptoms together, thus comprising a syndrome.

Many women experience mood disruptions or physical complaints during the late luteal phase. For this reason, the PMDD diagnosis has been controversial among psychiatrists, psychologists, gynecologists, and sociologists. One reason for the ambiguity may be the failure to distinguish normal from pathological premenstrual mood disturbance. Unlike other psychiatric syndromes, the remission and relapse of premenstrual symptoms is connected to a physiological process, the menstrual cycle. Every woman does not experience a pathological premenstrual mood disorder, however. Furthermore, there are discrepant views in defining depression as a categorical versus spectrum illness. Depression can be defined as a discrete, all-or-none variable or on a continuum of symptom severity. Thus, the differentiation between the normal and pathological is important in terms of investigative research, interdisciplinary discourse, as well as treatment.

Premenstrual criteria were included in the appendix of DSM-III-R as late luteal phase dysphoric disorder (LLPDD), and they are currently included as PMDD in the appendix of the DSM-IV-TR as a depressive disorder, not otherwise specified. Some researchers have postulated that moving PMDD into the depressive disorders not otherwise specified section of the DSM-IV has reinforced the view that PMDD is depression, rather than a unique disorder. Moreover, the treatment of PMDD with selective serotonin reuptake inhibitors (SSRIs) may be treating symptoms of depression rather than the etiological disturbance; thus, treatment may provide little information on the nature of the disorder and raises the problem of comorbidity with other mood disturbances (e.g., major depressive disorder).

Besides emotional distress, the disturbance must cause impairment in occupational or social functioning. Also, the mood disturbances must begin during the late luteal phase and remit at the follicular phase; this trend of mood fluctuation must occur over at least two consecutive cycles. Because retrospective ratings may not be accurate due to an expectation bias, premenstrual trends must be documented by daily mood ratings. In the United States, approximately 3–5% of women are thought to experience symptoms that meet criteria for PMDD. Even with the DSM-IV-established criteria, debate over defining criteria of PMDD is ongoing. The requirement of absence of symptoms postmenstrually may be unreasonably conservative, and the criterion for premenstrual interference in daily activities contributes little to the diagnosis of PMDD. Poor classification proves to be particularly problematic because, over time, untreated PMDD becomes progressively more severe, and episodes of dysphoria extend in duration. For this reason, women diagnosed with PMDD are vulnerable to developing major depressive disorder at a later date.

Etiology

Various disciplines conceptualize the etiological basis of premenstrual mood disturbance. Biological theorists maintain that premenstrual symptoms originate from the biological reactivity or sensitivity to reproductive hormones. Psychological theorists suggest that premenstrual symptoms arise from the maladaptive response of the individual to the environment. Social theorists assert that the social, patriarchal environment elicits symptoms. Regardless of the etiological model, hormonal changes associated with the menstrual cycle are likely to serve as contributory factors to mood disruption.

Biomedical Model

Investigators and clinicians have recognized the spectrum of premenstrual mood: symptoms ranging from mild to severe. Those with severe symptoms have a decreased ability to function effectively and meet criteria for PMDD. Rather than psychosomatic origins, researchers have investigated physiological underpinnings to the disorder that are expressed in mood and behavior, that is, conceptualized as a somato-psychic illness.

Estrogen and progesterone Research in this area began with the hypothesis that the absolute amount of estrogen (E2) and progesterone (P4), e.g., P4 deficiency, E2 excess, were responsible for mood disturbances. Inconsistencies in research findings have called this hypothesis into question. The rate of hormonal change, however, may play an important role in the pathogenesis of the disorder. Schmidt and colleagues examined the role of E2 and P4 by testing the effects of ovarian suppression with leuprolide, an agonist analog of gonadotropin-releasing hormone, versus placebo in women with and without PMS. The authors found that leuprolide caused a significant decrease in symptom scores for women with PMS compared with placebo treatment. In a subsequent phase of the study, Schmidt and colleagues administered a conjunct hormone therapy (leuprolide plus estradiol or progesterone) to the women with PMS whose symptoms responded to ovarian suppression, which resulted in a significant recurrence of symptoms. Thus, baseline levels of gonadal steroids or changes in hormone levels triggered a change in mood in women sensitive to mood instability. This study suggests that the occurrence of PMS symptoms may represent an abnormal response to normal hormonal changes. In addition, it suggests that E2 and P4 variations likely exert their influence on mood indirectly rather than directly. The central nervous system (CNS) is thus a nontraditional target for sex steroids.

Other neuroendocrine variables Research also suggests that abnormal fluctuations in thyroid, cortisol, prolactin, mineralocorticoids, prostaglandins (PGs), and endogenous opiates correspond to mood changes during the menstrual cycle. In addition, women with PMDD have a brief, transient, heightened cortisol response to ovine (o) corticotropin-releasing hormone (CRH) stimulation, blunted cortisol responses to serotonergic agonists in the late luteal phase, and lower evening plasma cortisol levels. Alterations in phase but not amplitude measures of cortisol have been found in women with PMDD during the

menstrual cycle. These findings suggest a physiological dysfunction between the hypothalamic-pituitary-adrenocortical (HPA) and the hypothalamic-pituitary-ovarian (HPO) axis for women with PMDD. This relationship may be complex and include a variety of other systems in its interaction.

Neurotransmitters Current research suggests that neurotransmitter abnormalities in PMDD are similar to those implicated in major depressive disorder. Specifically, PMDD patients have been differentiated from normal control patients by deficiencies in the serotonergic system. Investigators have found that women with PMDD have reduced platelet uptake of serotonin 1 week before menstruation, low whole blood serotonin during the last 10 days of the menstrual cycle, and an abnormal response to tryptophan loading in the late luteal phase. SSRIs (e.g., fluoxetine, sertraline, paroxetine, fluvoxamine) have been found to be efficacious in the treatment of PMDD. The efficacy of this pharmacological class is further evidenced by the FDA approval of fluoxetine and sertraline for PMDD. Recent work has shown that fluoxetine, citalopram, clomipramine, and sertraline can be effective if administered intermittently during the luteal phase. Further work is necessary to corroborate these findings. This medication regimen may be preferable to some women who do not wish to take chronic medication for a periodic condition. Women on a luteal dosing schedule generally reach response rates of 65%, with some women experiencing no significant benefits. This response rate may be reduced by nonadherence with a luteal phase dosing regimen.

In addition to serotonin, norepinephrine has been examined as another neurotransmitter that might contribute to PMDD. However, trials comparing fluoxetine to bupropion, sertraline to desipramine, and paroxetine to maprotiline demonstrate that augmenting noradrenergic activity alone is not effective for the treatment of PMDD.

Circadian rhythms Besides antidepressant medication, sleep and light therapies also have been found to be beneficial in reducing symptoms. The efficacy of these treatments is based on a chronobiological hypothesis of depression. Chronobiological investigations indicate that people with major depressive disorder may have disruptions in the internal regulation of rapid eye movement (REM) sleep propensity, temperature, cortisol, and melatonin, which occur throughout the 24-h daily cycle. Furthermore, the oscillator regulating these circadian rhythms may be desynchronized with the sleep–wake cycle. External cues, such as bright light, play an important role in the modulation and synchronization of these circadian rhythms. For example, exposure to 500 lux of light acutely suppresses melatonin secretion, a marker for circadian phase position. Light also shifts circadian rhythms and melatonin. Research indicates that patients with affective illness may be supersensitive to the suppressive effects of light on melatonin secretion, possibly due to inadequate exposure to daytime light; they may have altered amplitude of melatonin secretion and body temperature.

There may be differences between men and women in the regulation or expression of circadian rhythms. Compared with men, menstruating women have a shorter free-running period (i.e., length of time of a rhythm in an environment free of external cues, such as light), longer sleep duration, and lower amplitude in body temperature that varies with phases of the menstrual cycle. These findings suggest that reproductive hormones (including testosterone) affect circadian rhythm amplitude and synchronization. Indeed, comparative studies in rodents show that estrogen advances and progesterone delays circadian rhythms. Estrogen also appears to enhance synchronization between the oscillators.

Thus, fluctuations of menstrual cycle hormones may alter circadian rhythmicity. Instability of circadian rhythms resulting from changing reproductive hormonal levels may increase the risk for depression in some women. Chronobiological hypotheses contend that disrupted circadian rhythmicity is a factor contributing to depression. Women with PMDD have been reported to have decreased melatonin amplitudes, higher nocturnal body temperatures, and disturbances in the timing and amplitude of prolactin and thyroid-stimulating hormone (TSH) secretion. Also, the offset and duration of the melatonin rhythm may be disturbed and show an abnormal response to light. Light and sleep therapies may relieve PMDD symptoms by helping to realign underlying circadian clocks.

Psychological Model

In this model, women experiencing pathological premenstrual symptoms are distinguished from healthy women on the basis of the way they perceive or react to the environment. Therefore, treatment of PMDD is often directed toward physical activity, psychotherapy, relaxation training, support groups, and stress management.

One psychological model has studied PMDD from a state-dependent learning perspective, while biological markers tend to be trait dependent. Compared

with healthy subjects, women with menstrually related mood disorders report more negative life events and greater distress associated with those events only in the late luteal phase of the cycle. Thus, symptoms are conceptualized as occurring in an experiential state, such as the late luteal phase, which is characterized by affective and cognitive components. In support of the state-dependent model, women with severe premenstrual symptoms were more likely to use emotionally based coping techniques (i.e., catharsis) in the premenstrual phase than other strategies, such as cognitive reappraisal. Also during this phase, women with premenstrual dysphoria may be more likely than healthy controls to seek social support when perceiving daily stressors as uncontrollable.

Other findings fail to support the state-dependent hypothesis. Woods and colleagues suggested that a general stressful life context predicts premenstrual symptoms more than stress experienced during a particular menstrual cycle phase. A stressful milieu was a large predictor of perimenstrual symptoms and especially negative affect. Major life stressors play an important role in the development of depressive symptoms; however, the cumulative effect of daily stressors may be more influential in the generation of global premenstrual symptoms.

Type of stressor may also affect the manifestation of symptoms. For example, some research indicates that women with premenstrual dysphoria experience negative interpersonal interactions. Kuczmierczyk, Labrum, and Johnson found that women with PMDD perceive their work and especially their family lives as more stressful. Increased conflict and less self-sufficiency and assertiveness were noted in family environments of PMDD patients, which suggest that enmeshed and rigid family structures may have a role in the development of PMDD. These types of family structures have been previously implicated in other somatization processes.

In addition to disturbances in mood, women with PMDD may experience physical and cognitive changes. Cognitive deficits in major depression have been well documented and often include disturbances in attention, forgetfulness, psychomotor retardation, and proneness to confusion. In some cases, these cognitive disturbances may be secondary to mood disturbances. However, in the case of PMDD, the role reproductive hormone fluctuation may have on memory and cognitive functioning recently has been the focus of several studies. In healthy women without psychiatric illness, performance on neuropsychological tests has been found to vary with the monthly fluctuation in reproductive hormones. High levels of E2, as experienced in the late follicular phase, may elevate performance on automatization abilities, perceptual motor speed, mental arithmetic, and verbal memory. Lower levels of E2, as experienced during the menstrual phase, may be associated with higher performance on visuospatial tasks. The majority of these effects have been subtle and were not able to be replicated; this effect may be due to difficulties in documentation and measurement of ovulation, plasma hormone levels, mood symptoms, and individual variation in cycle length. Furthermore, cyclic performance changes are minor and are unlikely to interfere with normal functioning.

Social Model

From the earliest times, various facets of women's personalities, capabilities, and moods have been attributed to menstruation. Instability, resulting from women's reproductive cycles, has been used to justify denying women equal access to education and employment. One feminist model maintains that premenstrual syndrome pathologizes a naturally occurring phenomenon. Theorists state that high prevalence rates of premenstrual symptoms support this assumption and that the distinction of a separate pathological disorder is a spurious dichotomy. Based on historical comparisons of PMS and hysteria, writers have suggested that the use of a biological explanation for this disorder may legitimize societal prejudices. Furthermore, social theorists state that the medical model may pathologize normal biological processes and oversimplify women's feelings, problems, and conflicts through the diagnosis of PMDD. In fact, medical investigations have revealed that increases in cyclicity can be viewed as adaptive because it leads to homeostasis, stabilization, and longevity in life. Thus, it is the diagnosis of PMDD that leads to the medicalization of women's experiences of the menstrual cycle and not the cycle-related changes themselves. It is suggested that PMDD is a socially constructed disorder rather than a psychiatric disorder.

Since Western society is a contributory factor in this model, research on the prevalence of PMDD in other cultures is of importance in understanding the universality of symptoms. A limited number of studies have examined symptoms of PMDD in non-Western societies. This research has been plagued by a number of important methodological concerns and questionable validity of instruments that have been translated for use in populations other than those in which they were developed. In addition, the sociopolitical position of women in other cultures and their societal beliefs and expectations about menstruation are difficult to quantify and thus have not been taken

into account. Despite these difficulties, cross-cultural studies are of interest in considering the relative importance of biological versus sociological factors in the etiology of PMDD. Investigative studies have indicated a widespread variance in the type and severity of premenstrual symptoms experienced, which supports the interpretation of normal cyclical change. Eriksen found that women in developing countries tended to report shifts in mood associated with the onset of menstruation or throughout the menstrual cycle. Several studies have found reports of higher rates of somatic than affective symptoms in populations outside the United States, suggesting that culture may influence the symptoms that women notice, as well as the symptoms they determine to be problematic during the premenstrual period. As a result of these observed cross-cultural differences, menstrual socialization has been proposed as a determinant of premenstrual complaints. These findings suggest that PMDD is a culture-bound syndrome specific to Western cultures in which most women have been socialized to have negative expectations about menstruation. More specifically, it is argued that North American culture and media perpetuate the idea that the premenstrual period will be associated with negative affect and mood instability, causing women to interpret normal physiological changes that are essentially neutral in nature to have negative connotations.

Emergence of a Biopsychosocial Model

Since the relatively recent invention of radioimmunoassays, biomedical research has clearly dominated the area of PMDD. Because PMDD symptoms are associated with hormonal fluctuations, cultural and psychological phenomena associated with the disorder have often been ignored in scientific research, which has resulted in a poor integration of the findings. In the field of depression and other mood disorders, multidimensional theoretical models have emerged and contributed greatly to understanding the experience of depression for an individual. In addition, exploring the interaction of biological, psychological, and social processes has expanded the range of treatment options and provided individuals who suffer from depression effective alternatives to medication. For similar reasons, an integrative approach could greatly benefit research and clinical work in the area of premenstrual mood disturbances. A two-phase approach to encompass a comprehensive biopsychosocial perspective is recommended, with the first phase emphasizing self-monitoring through charting, stress management, and dietary and exercise programs. In addition, cognitive-behavioral coping strategies and reattribution training may help alleviate symptoms. The second phase consists of a pharmacological approach to mitigate biological aspects contributing to PMDD.

In the area of affective illness, a social zeitgeber theory of mood disturbance has been one unifying hypothesis linking biological and psychosocial models. Social zeitgebers are personal relationships, social demands, or tasks that entrain biological rhythms. Similar to the effect of light, social events are external regulators that synchronize circadian rhythms. For instance, marriage, birth of a child, divorce, and loss of a job are social events that disrupt natural mealtimes, sleeping times, and times of activity, which affect the stability of biological rhythms. In individuals predisposed to affective disruption, disturbances in the biological clock may develop into a state of ongoing desynchronization as observed in major depression. Social rhythm disruptions also have been implicated in the onset of manic episodes. Ehlers and colleagues suggested that both interpersonal psychotherapy and cognitive-behavioral therapy, two empirically supported treatments for depression, address the regularity of social routine. In addition, they state that the model is likely to be influenced by personality factors, gender, social support, coping, genetic/familial loading, and past treatment experience.

Adoption of the social zeitgeber model for premenstrual mood disturbances may potentially elucidate the interaction of important psychosocial and biological etiological variables. Socially, women may be particularly prone to respond more dramatically to interpersonal conflicts because they may rely on social zeitgebers, or social relationships, more than men. Some feminist models may explain this perspective. For instance, Carol Gilligan proposed that women may view relationships in terms of interconnections and caring, and men may view relationships in terms of hierarchy and power. Developmental research has documented that girls may be more likely to engage in prosocial behavior than boys. On self-report measures, women are often more empathetic and nurturing than men. The emphasis that women may place on connection with others possibly serves as a vulnerability to a mood disturbance; interpersonal disruptions may have a greater impact on women than on men. Animal and human studies have also shown that responses to social interactions can lead to fluctuations in whole blood serotonin levels. Therefore, women may be able to regulate serotonin function by seeking appropriately rewarding social interactions; the inability to seek or receive the social interactions could lead

to further physiological dysregulation via circadian rhythm disruption and/or reduction of serotonin, manifesting as symptoms of PMDD.

At the same time, this increased need for social support may be more adaptive in the long term; women may cope with stress, such as job loss or a death, better than men. Therefore, the increased variation in mood may be more adaptive by enhancing the mechanism of homeostasis, thus protecting women against other types of long-standing illness and increasing longevity.

These correlations between behavior and neuroendocrinology have also been explored within a framework of functional hypothalamic amenorrhea. For some time, researchers have noted that psychosocial variables, such as exercise, personality traits, and environmental stress, have the potential to induce ovarian acyclicity by inhibiting the release of gonadotropin-releasing hormone (GnRH) through mediating variables, such as CRH, endogenous opioids, dopamine, and TSH. Further work in this area potentially could elucidate important somato-psychic interactions, which are key to understanding the etiology of PMDD.

The biological, psychological, and social approaches to PMDD taken independently do not take into account the more complex relationships that formulate the disorder. Practically, it may be useful to examine PMDD through a single pathway that focuses on main effects and a linear relationship. However, lack of attention to the cluster of influences and their dynamic relations may oversimplify pathways as to how individuals develop severe premenstrual disturbances. Increasingly, the biopsychosocial model of PMDD is being examined to generate a more comprehensive model. Although great strides have been made in the past decade in increasing our understanding of PMDD, we must continue to supplement current approaches with research that focuses on more complex interactions among various factors. This future direction can only be accomplished through the continued interdisciplinary effort on a multivariate, biopsychosocial approach to PMDD.

See Also the Following Articles

Anxiety; Depression and Manic-Depressive Illness; Menstrual Cycles and Stress; Selective Serotonin Reuptake Inhibitors (SSRIs).

Further Reading

Blak, F., Salkovskis, P., Gath, D., et al. (1998). Cognitive therapy for premenstrual syndrome: a controlled study. *Journal of Psychosomatic Research* **45**(4), 307–318.

Dimmock, P. W., Wyatt, K. M., Jones, P. W., et al. (2000). Efficacy of selective serotonin-reuptake inhibitors in premenstrual syndrome: a systematic review. *Lancet (North American Edition)* **356**, 1131–1136.

Endicott, J. (2000). History, evaluation and diagnosis of premenstrual dysphoric disorder. *Jounral of Clinical Psychiatry* **61**(12), 5–8.

Eriksson, E., Andersch, B., Ho, H. P., et al. (2002). Diagnosis and treatment of premenstrual dysphoria. *Journal of Clinical Psychiatry* **63**(7), 16–23.

Halbreich, U., Bergeron, R., Yonkers, K. A., et al. (2002). Efficacy of intermittent, luteal phase sertraline treatment of premenstrual dysphoric disorder. *Obstetrics & Gynecology* **100**(6), 1219–1229.

Parry, B. L. (2001). The role of central serotonergic dysfunction in the aetiology of premenstrual dysphoric disorder: therapeutic implications. *CNS Drugs* **15**(4), 277–285.

Parry, B. L. and Berga, S. L. (2002). Premenstrual dysphoric disorder. In: Arnold, A. P., Etgen, A. M., Fahrbach, S. E. & Rubin, R. T. (eds.) *Hormones, brain, and behavior* (vol. 5, chap. 99) pp. 531–552. New York: Academic Press/Elsevier Science.

Parry, B. L., Berga, S. L., Mostofi, N., et al. (1989). Morning vs. evening light treatment of late luteal phase dysphoric disorder. *American Journal of Psychiatry* **146**, 1215–1217.

Ross, L. E. and Steiner, M. (2003). A biopsychosocial approach to premenstrual dysphoric disorder. *Psychiatric Clinics of North America* **26**, 259–546.

Schmidt, P. J., Nieman, L. K., Danaceau, M. A., et al. (1998). Differential behavioral effects of gonadal steroids in women with and in those without premenstrual syndrome. *The New England Journal of Medicine* **338**, 209–216.

Pre-pulse Inhibition

M van den Buuse
Mental Health Research Institute, Melbourne, Australia

© 2007 Elsevier Inc. All rights reserved.

What Is Prepulse Inhibition?
Clinical Correlates
Effect of Stress on Prepulse Inhibition in Rats and Mice
Mechanism of Stress Effect on Prepulse Inhibition
Conclusion

Glossary

Prepulse inhibition (PPI)	A measure of sensorimotor gating, usually the reduction of the startle response to an intense stimulus if this is preceded by a low-intensity prestimulus.
Sensorimotor gating	A normal protective mechanism in the brain that functions to gate or filter irrelevant sensory stimulation, allowing for coherent thought.

What Is Prepulse Inhibition?

PPI is a measure of sensorimotor gating and is widely used to study information processing and attention dysfunctions. Sensorimotor gating is a normal protective mechanism in the brain that functions to gate or filter irrelevant sensory stimulation, allowing for coherent thought. Deficits in sensorimotor gating may result in stimulus overload and the misinterpretation of sensory information.

PPI of the acoustic startle response can be easily measured in several species, including rats, mice, and humans. The startle response is elicited by the presentation of short, loud acoustic stimuli (pulses or clicks) that usually induce whole-body startle responses in rats and mice and reflex eyeblink responses in humans. When the acoustic startle stimulus is preceded by a lower-intensity prestimulus (a prepulse), the subsequent behavioral response is reduced (see Figure 1). This mechanism is mediated by a well-defined neural circuit, including a short brain-stem pathway mediating the startle response and several modulating brain regions, such as the hippocampus, nucleus accumbens, and frontal cortex. PPI is deficient in a number of mental illnesses, including schizophrenia, obsessive-compulsive disorder, and attention-deficit disorder. In experimental animals, deficits in PPI can be induced by treatment with dopamine receptor agonists, serotonin 1A and 2A receptor agonists, or NMDA receptor antagonists. Several studies have shown that deficits in humans or treated animals can be reversed by treatment with antipsychotic drugs, such as clozapine. Because of the importance of stress as a risk factor in mental illnesses, studies have addressed the effect of stress on PPI.

Clinical Correlates

A clear example of a link between stress and PPI comes from observations in patients with posttraumatic stress disorder (PTSD). There was a significant reduction of PPI when 21 Vietnam veterans with PTSD were compared to a civilian control group of 17 subjects. However, when the 21-subject group was compared to an 11-subject Vietnam veteran control group without PTSD, the difference reached only trend level. The magnitude of PPI disruption was correlated with the extent of PTSD symptoms (Mississippi scale), specifically the extent of intrusive memories and the level of arousal. Subsequent studies in other patient groups by the same authors have only partially replicated these findings. Similarly, whereas in a groups of 8- to 13-year-old children with PTSD PPI was clearly diminished, in another study of a group of 28 adolescent girls with PTSD, there were no significant differences in PPI compared to a group of 23 healthy nontraumatized girls. These discrepancies between studies could be related to the differences in the causes of PTSD in these subjects and also to parametric and methodological differences between the studies.

Many PPI studies have focused on schizophrenia. The majority of these studies have found significant disruptions of PPI in these patients, which was reversed by treatment with antipsychotic drugs. However, although stress is recognized as a risk factor in this illness, few studies have investigated stress modulation of PPI in patients with schizophrenia (see **Schizophrenia**).

Effect of Stress on Prepulse Inhibition in Rats and Mice

In contrast to the paucity of clinical studies on stress and PPI, several experimental animal studies have addressed this relationship. These studies can be divided according to the timing of the stress: perinatal, chronic postweaning, or acute stress.

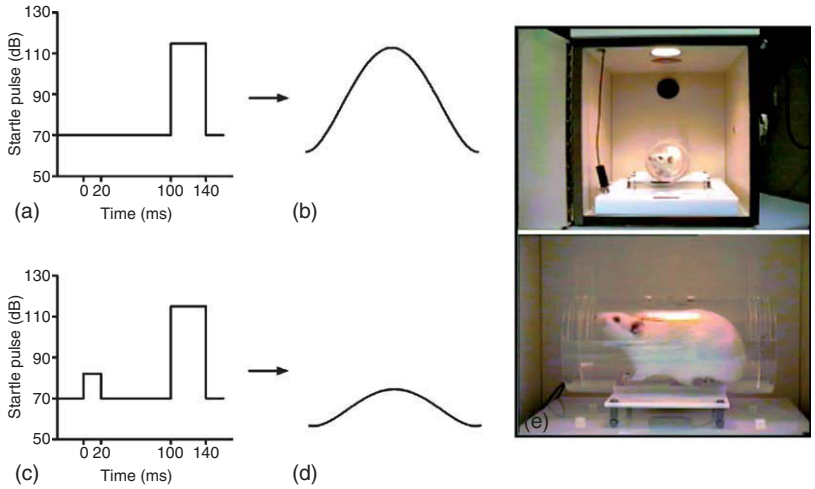


Figure 1 Prepulse inhibition of acoustic startle in rats. a, 40-ms 115-dB startle pulse; b, response in pulse-alone trials; c, startle pulse preceded 100 ms before by a 20-ms prepulse of 12 dB above baseline; d, reduced startle responses in a prepulse–pulse trial; e, rats in the automated startle apparatus used for PPI experiments. If PPI is disrupted because of pharmacological treatment or disease, the presentation of the prepulse–pulse trial evokes responses closer to or similar to those after pulse-alone trials.

Perinatal Stress and Prepulse Inhibition

The neurodevelopmental origin of schizophrenia has led to several studies on the effect of early stress on PPI. The rationale for these studies is that an early stressful event elicits subtle changes in brain development that result in changes in behavior comparable to symptoms of schizophrenia. When rats are born, they are much less developed than humans. Thus, neonatal maternal deprivation in rats has been used to study the effects of prenatal stress factors in humans. Several neuroendocrine studies have shown prolonged alterations in responsiveness of the hypothalamic-pituitary-adrenal axis in maternally deprived rats. A single 24-h separation of rat pups from their mother also induced a disruption of PPI later in life. This disruption developed only in adulthood, and the effect was strongly dependent on the timing of the stress. Treatment with typical or atypical antipsychotic drugs could reverse the disruption of PPI in these animals. Maternal deprivation did not cause PPI disruption in all rat strains tested, suggesting an important genetic modulation of stress sensitivity and PPI. Although baseline dopamine release in the forebrain nucleus accumbens was not altered, the stimulation of dopamine release by K^+ infusion or amphetamine treatment resulted in much greater increases in dopamine levels in maternally deprived rats than in controls.

The role of maternal care has also been identified. Rat dams may be grouped according to their maternal behavior into high- and low-licking/grooming dams. The offspring of low-licking/grooming dams, which are therefore subjected to relatively low levels of maternal care, showed lower PPI than the offspring of high-licking/grooming dams, which are subjected to relatively high levels of maternal care. This result again suggests that stress very early on in development can have long-lasting effects on behavior in general and on PPI in particular.

There are, furthermore, other studies using early stressors, such as neonatal endotoxin treatment in rats to simulate bacterial infection, or chronic stress in the pregnant dam. In many of these studies, the offspring develop deficits in PPI when reaching adulthood.

Postweaning Stress and Prepulse Inhibition

Isolation rearing in rats consists of housing the animals in single cages for a prolonged period of time. This procedure induces a significant disruption of PPI, although a comparison of different rat strains showed that this was not seen in all strains. The timing of isolation housing was also important, with the greatest disruption of PPI seen in animals that had been singly housed from weaning. Isolation rearing, and its associated behavioral perturbations, in rats has been used as a developmental animal model of schizophrenia. Indeed, disruptions of PPI in isolation-reared rats could be reversed by treatment with various antipsychotic drugs. Social isolation was associated with alterations in the release and turnover of dopamine and other neurotransmitters in different parts of the brain, which could explain the effect of antipsychotic drugs. For example, similar to maternally deprived rats, in isolation-housed rats amphetamine-induced dopamine release in the nucleus accumbens was markedly enhanced when compared to group-housed controls. A similar finding has been made in

patients with schizophrenia when compared to normal controls.

Acute Stress and Prepulse Inhibition

Only few studies have specifically addressed the effect of acute stress on PPI. It is important to note that the PPI procedure itself is mildly stressful in rodents. Acute treatment with antipsychotic drugs has been shown to result in a subtle but significant increase in PPI. This may be explained by these drugs blocking enhanced activity of dopaminergic projections and/or other neurotransmitter pathways in response to the stress of the PPI procedure. This activation can be exacerbated by acute stress shortly before the procedure, although such effects are seen only under certain conditions. For example, we have observed that the mild stress of a saline injection caused a disruption of PPI in some mice strains but not others. Similarly, injection stress caused a disruption of PPI in rats that had undergone neonatal depletion of dopamine but not in controls. Interestingly, in healthy human volunteers, mental arithmetic stress was shown to cause a disruption of P50 gating, another measure of sensory information processing. The disruption took P50 gating to the reduced level seen in patients with schizophrenia.

Mechanism of Stress Effect on Prepulse Inhibition

Peripheral Effects

Two of the hallmarks of the stress response are enhanced release from the adrenals of adrenaline and cortisol (corticosterone in rodents). In a number of studies, it has been shown that adrenalectomy does not significantly affect PPI. However, we found that the 25% disruption of PPI caused by treatment with the dopamine-releasing drug amphetamine was absent in adrenalectomized rats (Figure 2). The disruption of PPI caused by treatment with the dopamine receptor agonist apomorphine, the glutamate receptor antagonist MK-801 (Figure 2), and the serotonin-1A receptor agonist 8-OH-DPAT was not altered in adrenalectomized rats. This suggests a modulatory action of circulating factors from the adrenal, presumably corticosterone, on the brain mechanisms involved in the effect of amphetamine on PPI. These brain mechanisms probably include dopaminergic activity. This was also seen in other studies. Thus, in adrenalectomized mice, haloperidol treatment did not cause the increase in PPI seen in intact mice or adrenalectomized mice supplemented with moderate doses of corticosterone. Interestingly, the haloperidol-induced increase of PPI was also absent in mice treated with high doses of corticosterone. Similar results were seen in rats. Chronic treatment with relatively high doses of corticosterone caused a decrease of baseline PPI in rats. Apomorphine-induced disruption of PPI was markedly greater in these animals when compared to either intact controls or adrenalectomized rats (Figure 3). Taken together, these results could indicate that the dopaminergic modulation of PPI, here exposed by the action of haloperidol and of apomorphine, is dependent on certain optimal circulating levels of corticosterone. Any deviation from these optimal levels, either by adrenalectomy or treatment with high doses of the steroid, affected this mechanism.

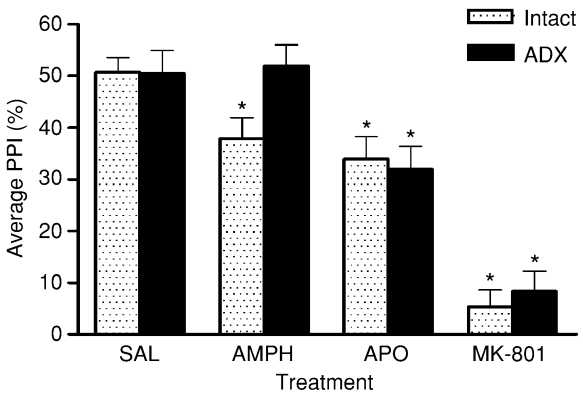


Figure 2 Average PPI of intact and adrenalectomized (ADX) rats after treatment with saline (SAL), 0.5 mg/kg of the dopamine-releasing drug amphetamine (AMPH), 0.1 mg/kg of the dopamine receptor agonist apomorphine (APO), and 0.1 mg/kg of the N-methyl-D-aspartate receptor antagonist MK-801. All drug treatments caused a disruption of PPI compared to saline treatment, except amphetamine treatment in adrenalectomized rats, which was ineffective. *$p < 0.05$ for difference with saline treatment in the same group. Based on unpublished data from Adams and van den Buuse; for methodological details, see van den Buuse et al. (2004).

Central Nervous System Effects

In the hypothalamus, corticotropin releasing hormone (CRH) and vasopressin are involved in the regulation of adrenocorticotropic hormone (ACTH) release from the pituitary. The central administration of CRH caused a disruption of PPI in rats. Similarly, in transgenic mice overexpressing CRH, there was a decrease of PPI that was reversed by treatment with haloperidol and risperidone. CRH 1 receptor (CRHR1) and CRHR2 appear to play opposing roles in mediating the effect of CRH on PPI. Thus, treatment with CRHR1 and CRHR2 antagonists caused an increase or a decrease of PPI, respectively. Furthermore, the disruption of PPI caused by CRH treatment could be blocked by pretreatment with a CRHR1 antagonist. In CRHR1 knockout mice, CRH did not disrupt PPI but, rather, enhanced it. Overall, these results could indicate that CRH in the brain plays an

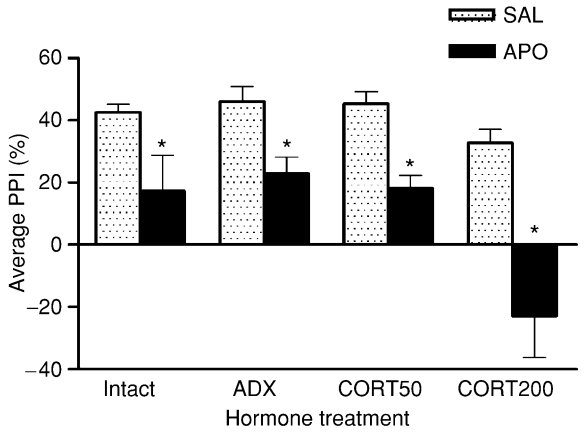


Figure 3 Average PPI of intact rats and of adrenalectomized rats that were implanted with either a cholesterol control pellet (ADX), a 50-mg corticosterone pellet (CORT50) or a 200-mg corticosterone pellet (CORT200). Two weeks after pellet implantation, the rats were injected with saline (SAL) or 0.1 mg/kg of the dopamine receptor agonist apomorphine (APO) and were tested for PPI. The 50–60% disruption of PPI observed in all groups was markedly enhanced (170%) in rats treated with a high dose of corticosterone. $^*p < 0.05$ for difference with saline-treatment in the same group. Based on unpublished results from Adams and van den Buuse; for methodological details, see van den Buuse et al. (2004).

important role in the effects of stress on sensorimotor gating. Conversely, drugs acting on CRH receptors could be relevant in reversing PPI deficits in patients with PTSD or schizophrenia.

There is limited information on the role of vasopressin in PPI. Vasopressin-deficient Brattleboro rats display disruptions of PPI that can be reversed by chronic, but not acute, treatment with haloperidol. Also vasopressin V1b receptor knockout mice show a disruption of PPI. Chronic treatment with the atypical antipsychotic drugs, risperidone and clozapine (but not the typical antipsychotic, haloperidol) reversed this PPI disruption. The authors noted decreased extracellular levels of dopamine in the frontal cortex of these animals; however, the activity of the hypothalamic-pituitary-adrenal axis was not investigated.

Conclusion

Studies on the modulation of PPI by stress and peptides involved in the stress response provide important insights into the role of stress in the symptoms of illnesses such as PTSD and schizophrenia. The effect of stress on PPI is mediated by circulating factors, such as corticosteroids, or central factors, such as CRH.

See Also the Following Articles

Corticosteroids and Stress; Corticotropin Releasing Factor (CRF); Dopamine, Central; Schizophrenia; Serotonin.

Further Reading

Ellenbroek, B. A. (2003). Animal models in the genomic era: possibilities and limitations with special emphasis on schizophrenia. *Behavioural Pharmacology* **14**, 409–417.

Geyer, M. A. and Markou, A. (1995). Animal models of psychiatric disorders. In: Bloom, F. E. & Kupfer, D. J. (eds.) *Psychopharmacology: the fourth generation of progress*, pp. 787–798. New York: Raven Press.

Koch, M. and Fendt, M. (2003). Startle response modulation as a behavioral tool in neuropharmacology. *Current Neuropharmacology* **1**, 175–185.

Swerdlow, N. R., Geyer, M. A. and Braff, D. (2001). Neural circuit regulation of prepulse inhibition of startle in the rat: current knowledge and future challenges. *Psychopharmacology* **156**, 194–215.

van den Buuse, M., Garner, B., Gogos, A., et al. (2005). Importance of animal models in schizophrenia research. *Australian and New Zealand Journal of Psychiatry* **39**, 550–557.

van den Buuse, M., Garner, B. and Koch, M. (2003). Neurodevelopmental animal models of schizophrenia: effects on prepulse inhibition. *Current Molecular Medicine* **3**, 459–471.

van den Buuse, M., Morris, M., Chavez, C., et al. (2004). Effect of adrenalectomy and corticosterone replacement on prepulse inhibition and locomotor activity in mice. *British Journal of Pharmacology* **142**, 543–550.

Zhang, T. Y., Chretien, P., Meaney, M. J., et al. (2005). Influence of naturally occurring variations in maternal care on prepulse inhibition of acoustic startle and the medial prefrontal cortical dopamine response to stress in adult rats. *Journal of Neuroscience* **25**, 1493–1502.

Pressor Responses *See:* Regional Blood Flow, Stress Effects.

Pressure, Effects of Extreme High and Low

R J Værnes
NUTEC Crisis Management, Ytre Lakevaag, and University of Bergen, Bergen, Norway

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R J Værnes, volume 3, pp 208–220, © 2000, Elsevier Inc.

Introduction
Humans Working under High Pressure
Humans Working under Low Pressure
Concluding Remarks

Glossary

Acute mountain sickness (AMS)	A high-altitude illness, which can be classified as normal or malignant AMS. Malignant AMS includes features of both pulmonary and cerebral edema.
Chronic mountain sickness (CMS)	A high-altitude illness whose symptoms are vague neuropsychological complaints, including headache, dizziness, fatigue, and difficulty concentrating.
High-pressure nervous syndrome	Behavioral disturbances after being at depths greater than 150 m. Central nervous system manifestations include tremor, dizziness, nausea, and electroencephalogram changes. It becomes more severe with increasing depth and during fast rates of compression.
Hypoxia	Lack of oxygen that results from, at normal pressure, a reduction of pO_2 and, at high altitude, the reduction in atmospheric pressure. Central nervous system manifestations differ between acute hypoxia (seconds to 1–2 h) and chronic hypoxia (hours to years).
Nitrogen narcosis	An illness from breathing compressed air at depths deeper than 20 m; characterized by euphoria, impaired cognitive functioning, and impaired neuromuscular coordination.
Oxygen toxicity	An illness that increases progressively with the increase of inspired pO_2 and with greater duration of exposure; central nervous system manifestations range from localized muscle twitching to grand mal seizures.
Residual effects	Residual impairment of central nervous system functioning following the return to normal pressure from deep depths and from high altitudes. For both high- and low-pressure conditions, the residual effects consist of neuropsychological complaints regarding cognitive functioning and fine visuomotor tempo and coordination.

Introduction

What Constitutes Extreme High- and Low-Pressure Conditions for Humans?

As we have throughout the centuries, today humans go into the sea for food, military operations, sport, treasure-hunting, and exploration. The earliest indications of breath-hold diving and the first techniques used by humans at work in the sea are from an archeological dig in Mesopotamia dating from 4500 BC and from Aristotle, who in the fourth century BC mentioned a type of diving machine. The historical stages in the development of diving are:

1. Breath-hold diving (also called free diving)
2. Diving bells
3. Surface-supported or helmet diving (also called hard-hat diving)
4. Self-contained underwater breathing apparatus (scuba diving)
5. Saturation diving

Diving physiologists had, by the late 1950s, made great strides toward understanding and managing some of the principal hazards that beset divers in their progress toward the depths. These hazards included oxygen toxicity, the narcotic effects of inert gases used as oxygen diluents in high-pressure environments, and the mechanical effects of dense atmospheres on respiratory work and the effectiveness of ventilation.

When oil and gas exploration started in the North Sea in the 1960s, diving became important for the maintenance, inspection, and repair of subsea installations. Therefore, the saturation diving technique was developed further, and diving to depths between 50 and 180 msw (meters of seawater) became a standard procedure. In parallel with this offshore diving, experimental studies on diving down to a maximum of 850 msw were conducted in the United States, Europe, and Japan. This research led to the observation of the high-pressure nervous syndrome (HPNS), a syndrome consisting of tremors, dizziness and nausea, electrophysiological changes, and impaired performance.

Humans, however, expose themselves to more than just extremely high pressures. For centuries, humans

have attempted to climb higher and higher, and scientists have been astonished repeatedly by mountaineers' ability to ascend higher than their confident predictions. Whereas the central nervous system (CNS) problems from diving are due to oxygen toxicity, inert gas narcosis, and HPNS, the CNS problems of extreme altitude are due to hypoxia (because there is reduction in atmospheric, and hence oxygen, pressure with increases in altitude) and cold (because of the temperature decrease with altitude and wind chill).

One of the earliest European reports of the problems of high altitude was given by the Greeks who, during their yearly ceremonial visit to the summit of Mountain Olympus (2911 m), were not able to survive unless they applied moist sponges to their noses (possibly soaked in vinegar). The first documented description of mountain sickness comes from Chinese sources in 32–37 BC. It has been suggested that the occurrence of mountain sickness and its complications may have been taken by the Chinese as a sign for them not to transgress their natural boundaries.

In the late nineteenth and early twentieth centuries, there was increasing scientific and medical interest in both the oxygen transport system and altitude adaptation, and many discoveries were made. Viault recorded the best-known physiological response to high altitude – an increase in the number of red blood cells. The tempo of Himalayan exploration accelerated during this period, and in 1892 while on Pioneer Peak (7000 m), Conway described the debilitating effects of long periods at high altitude. In 1953, the first ascent of Mt. Everest (8848 m) was conducted by Hillary and Tensing. However, they used supplementary oxygen. In 1978, Messner and Habeler made the first ascent of Mt. Everest without supplementary oxygen, and in 1987 the same was done in the winter (December) by Sherpa Ang Rita.

What Constitutes the Physical Characteristics of These Conditions?

Diving Awareness of the natural laws covering diving, and the results of disobedience, has fostered the development of the systems and techniques that allow humans to dive safely to increasingly greater depths The laws are the ideal gas law (behavior of gases at low pressures), Boyle's law (volume is inversely proportional to pressure), Charles's law (volume is directly proportional to temperature), real gas law (behavior of gases at high pressures), Dalton's law (law of additive partial pressures), Graham's law (diffusion rate is proportional to molecular rate), Amagat's law (law of additive partial volumes), thermal laws (specific heat, conduction and convection, and heat transfer), acoustical laws (conduction, diffraction, and the speed of sound), density laws (Archimedes's principle), and optical laws (curvature of light and absorption of color). A direct line can be drawn from any existing diving technique to at least one of these fundamental laws.

When people dive down to a certain depth, a compression profile is followed. For each 10 msw of depth, the pressure increases 1 atm absolute (ATA). Especially in the 1970s, surface-oriented diving with a wet diving bell was common practice. Two men descended, for example, in the bell to approximately 100 msw (11 ATA). The compression rate was normally in the range of 30 msw min^{-1} (3 ATA min^{-1}). The average time at maximum pressure was less than 30 min, and the tasks involved light or moderate effort. Only one diver exited to the work site while the other monitored his or her activity and stood by to assist.

Because the time required to come back to surface (decompress) is often longer than the time spent on the bottom, the ratio of diver work time to decompression time is not always a favorable one (if a diver decompresses too fast, he may get the bends, that is, gas bubbles in the skin, joints, spinal cord, or bloodstream).

The price that a diver must pay for staying deeper and longer led to the development of saturation diving. First accomplished in 1938, this diving technique was developed further in parallel with the oil and gas exploration in the North Sea. If a team of divers are to work at 150 msw (16 ATA), they compress to 16 ATA in a living chamber at a diving platform or vessel and stay at that pressure for 10 days. Pairs of divers then enter a dry diving bell, which is lowered to the work site at 150 msw. The bell stays at that depth for up to 8 h, and one diver at a time works in the water for half the period. After the 8-h period, the bell is taken up to surface and connected to the living chamber where the divers eat, relax, and sleep (all the time under the same pressure). After 10 days, they start decompression to surface pressure, which takes several days.

High altitudes In 1878, it was known already that most of the deleterious effects of high altitude in humans were caused by hypoxia, a direct result of the reduction in atmospheric pressure. However, there is still confusion about the relationship between barometric pressure and altitude, particularly at extreme heights. For example, the barometric pressure at the summit of Mt. Everest is considerably higher than that predicted by the standard pressure–altitude tables used in aviation.

Although most of the high-altitude effects are due to hypoxia, additional deterioration results from cold, dehydration, and solar and ionizing radiation.

Most of these additional hazards, however, can be avoided by using proper clothing or shelter. Only hypoxia is unavoidable unless supplementary oxygen is available. Low barometric pressure itself has no physiological consequences unless the decompression is rapid. For example, when a window fails in an aircraft, an explosive decompression can occur; and, just as in diving, rapid decompression causes barotraumas as a result of the very rapid enlargement of airspaces within the body, especially in the lungs and middle ear cavity.

Humans Working under High Pressure

Working Conditions (Types of Diving)

Helmet diving The first major breakthrough in helmet diving was the invention of an open diving dress in 1819 by Siebe, a waist-length jacket, a metal helmet, and air under pressure supplied from the surface to the helmet. Siebe later modified his open dress to the closed type, retaining the helmet and force pump and enclosing divers in a continuous airtight dress. Air was vented by a valve, the basis for modern hard-hat diving gear.

As advances were made in helmet diving, divers gained more air, more protection, and more mobility, but it was not until the development of self-contained underwater breathing apparatus (scuba) gear that they were afforded the freedom that breath-hold divers had to move about, coupled with an air supply that enabled them to stay down longer than free divers.

Scuba diving The Cousteau–Gagnan valve (aqua lung, presented in 1943) is the basis for modern open-circuit scuba gear. Open-circuit breathing gear usually contains compressed air as a breathing mix. It has a demand regulator, which provides air when the diver inhales, and the air is vented into the water on exhalation. To conserve the breathing mix and extend the time at depth, closed-circuit and semiclosed rebreathers have been developed.

A closed-circuit apparatus can employ pure oxygen or a mixed gas with oxygen and a dilutent gas such as nitrogen or helium. The risks inherent in breathing pure oxygen, primarily oxygen toxicity, limit the operating depths to approximately 10 msw or 18 msw resting.

When divers go deeper than 30 msw on air, dangerous signs and symptoms of euphoria, intoxication, and narcosis occur. The first inference that the nitrogen constituent of compressed air is responsible for this narcosis was made by Behnke and colleagues in 1935. They concluded that compressed air at depths deeper than 20 msw exerts a narcosis characterized by "euphoria, retardation of the higher mental processes, and impaired neuromuscular coordination" (Behnke et al. 1935). The signs and symptoms of the narcosis are similar to those due to drinking alcohol, hypoxia, and the early stages of anesthesia. Exposure to depths greater than 90 msw may also result in loss of consciousness.

Saturation diving The use of saturation diving in a commercial application was, in fact, suggested by Behnke as early as 1942. He noted that frequent decompressions were "not only potentially dangerous, but highly uneconomical. It would appear advisable therefore to keep men at work on a job continually under pressure" (Behnke 1942).

Although the first saturation dives were made using air, this method has never been used for diving at a depth less than 40 msw. For diving deeper, however, this has been the main diving technique since the early 1980s. What they initially had to do, however, was remove the nitrogen (which would lead to narcosis) and use helium instead. The breathing gas in saturation diving is therefore called heliox (helium and oxygen).

Modern-day saturation diving platforms in the North Sea have greatly extended diving capabilities. Up to 28 people can be in saturation, making excursions to different depths with two bells. Compression to depths greater than 150 msw, however, produces the signs and symptoms of high-pressure nervous syndrome (HPNS). The syndrome includes tremor in the hands and arms, electroencephalograph (EEG) changes, dizziness, nausea, and vomiting. At depths greater than 300 msw, lapses of consciousness may occur. The symptoms of HPNS become more severe with increasing depth and during fast rates of compression.

Effects on the Central Nervous System

As already indicated, several CNS reactions may occur when divers are exposed to this high-pressure environment. The stress reaction in the individual diver, however, is a product of a combination of factors, psychological and physiological, and is often a synergistic effect .This section focuses on the following six factors, which both alone and in combination can lead to severe stress and, in some cases, a fatal outcome in diving: (1) anxiety and panic, (2) oxygen toxicity, (3) nitrogen narcosis, (4) HPNS, (5) isolation and confinement, and (6) residual effects.

Anxiety and panic Stress during diving is probably the central problem causing the accidents and resulting injuries and fatalities that occur in sport divers.

Most diving-accident researchers implicate panic, as a response to stress, as the major cause of diving fatalities. The skills learned in training, such as the careful use of equipment and the techniques of coping with accidents, appear to be lost when divers panic. Therefore, it is important for divers, especially sports divers, to understand stress and to learn to handle stress.

Divers' use of defensive reactions to a threat also lead to inadequate performance under stress, and decrements in both defense mechanisms and performance relate to endocrine stress reaction. The first study demonstrating that perceptual defense mechanisms, as tested by the defense mechanism test (DMT), predicted individual reactions to dangerous situations was on attack divers. Three groups of divers ($N = 29$, 20, and 29) were tested with DMT, and the protocol was coded, rated, and ranked independently by three raters. All three raters' predictions, based exclusively on the DMT results, correlated significantly with the pass/fail criteria at the school.

In later studies in which more specific performance criteria were used, the same relationships were found. In a study of 45 navy divers, the individual score on defense mechanisms correlated significantly with the reduction in reasoning capacity at 60 msw. In a study of 29 navy divers, the same relationship between defense mechanisms and mental performance was found. In addition, the factor pattern of the endocrine postdive measures showed cortisol, catecholamine, and testosterone factors. There was a correlation between the cortisol factor and both defense pattern and performance decrement, confirming previous studies from other high-risk environments (parachutists).

Oxygen toxicity The greater application of hyperoxygenation both in therapy and in diving since the early 1970s has led to a better understanding of the universal dependence of vital biological processes on the cellular antioxidant defense mechanisms that have evolved as an adaptation to Earth's atmospheric oxygen tension. It is now apparent that the same oxygen pressures required to sustain life cause lethal oxygen poisoning in the absence of these mechanisms.

The severity of oxygen poisoning increases progressively with the increase in the inspired partial pressure of oxygen and greater duration of exposure. At sufficient pressure and exposure duration (i.e., an attack diver breathing pure oxygen at too great a depth), oxygen initially causes functional impairment and ultimately causes the chemical destruction of any living cell.

Donald exposed divers to oxygen pressures of 3 ATA or higher until CNS effects were experienced. Unfortunately, the sequence of progression did not consistently include minor symptoms before the onset of convulsions. Even when preconvulsive aura was experienced, it was sometimes followed so rapidly by seizures that it had little practical value as a warning. An enormous variation in oxygen tolerance was found among divers exposed to the same conditions. Attempts to correlate CNS oxygen tolerance with factors such as age, weight, physical fitness, smoking, alcohol ingestion, psychological stability, or personality traits were all unsuccessful. Even in the same diver, there was considerable variability in oxygen tolerance from day to day. The conclusion was that variability in oxygen tolerance from day to day within the same subject was as great as between subjects.

The seizures caused by oxygen toxicity are essentially the same as those occurring in grand mal epilepsy. They can be preceded by an aura or by a sequence of premonitory sensations, or they can occur suddenly. In contrast to the cerebral hypoxia that occurs during an epileptic seizure, brain hyperoxygenation is maintained during an oxygen convulsion by a high alveolar partial pressure of oxygen in conjunction with marked hypercapnia and increased cerebral blood flow. Biochemical studies on free radicals under hyperoxic conditions have shown, however, that extracellular hyperoxia leads to cellular hypoxia through free-radical mechanisms.

However, it is not only the use of pure oxygen in diving that can lead to serious CNS problems. Ongoing research has shown that there is probably a synergistic effect between oxygen (in heliox with a partial pressure of O_2 of only 0.6 ATA) and high pressure when people dive to extreme depths (>300 msw), contributing to the explanation of CNS-related symptoms and signs (e.g., HPNS) that are observed.

In a series of experiments, the effects of various partial pressures of oxygen (0.2, 0.4, and 0.6 ATA) in the heliox breathing gas on HPNS were studied in rats. Rats were exposed to 91 ATA, equivalent to approximately 45 ATA for humans. The results showed that HPNS symptoms were reduced significantly with a lowered oxygen partial pressure. Follow-up studies on the effects of hyperbaric oxygen (>95%) on neuronal transmission in the perforant path/dentate area of hippocampal formation showed that epileptic seizure activity produces long-term changes in synaptic excitability in granular cells of the dentale area.

The inhibition of nitric oxide (NO) production has been reported to delay the onset of hyperbaric oxygen-induced epileptic seizures in mice and in rats. The mechanism(s) behind this protective effect is still equivocal. It has been suggested that NO is

involved in the regulation of blood flow and arterial pressure, as a transmitter in peripheral nerves, and as an intercellular or, rather, transcellular transmitter between CNS neurons. The protective effect of inhibiting NO production could, therefore, be due to changes in cerebral blood flow, neuronal activity, or both.

Nitrogen narcosis All divers know that diving deeper than 30 msw on air can lead to dangerous symptoms and signs of euphoria, intoxication, and narcosis. One of the first reports of nitrogen narcosis was given by a Frenchman in 1835. Around the turn of the twentieth century, it was reported by several others in tunnel workers breathing air at 46 m. However, as mentioned previously, Behnke and coworkers were the first to claim that nitrogen was the agent responsible for the narcosis (Behnke 1942).

As might be expected, there is a wide variation in individual susceptibility. Whereas one diver becomes narcotic at pressures as high as 50 msw, another is affected quite severely at only 30 msw. In this connection, Cousteau claimed that emotionally stable individuals are usually affected to a smaller extent than less stable subjects. This was confirmed later in experimental studies by Værnes in which divers sat in a pressure chamber taking cognitive tests on paper and performing no other physical activity.

In both studies, a significant reduction in long-term memory (LTM) was found. When information input was done predive and recall at 60 msw, the mean reduction in LTM was 31.2%. When input was done at 60 msw and recall later at the surface, LTM impairment was even worse (45.6%). (The control situation was both input and recall at the surface in the same chamber with the same duration between input and recall.) Memory impairment was obviously a major effect of nitrogen narcosis. As reported earlier, some divers also had a reduction in reasoning, although the average reduction was not significant. However, for reasoning impairment, in both studies there was a significant correlation with the degree of defensive reaction pattern as defined by DMT, and in the second study, it also correlated with the degree of cortisol activation.

High-pressure nervous syndrome Compression in a heliox atmosphere induces a number of behavioral disturbances at depths greater than 150 msw (16 ATA). These disturbances were described by Bennett as nausea, vertigo, and tremor during rapid compression to 180 msw, and he attributed them to the helium (i.e., helium tremor). Brauer and colleagues and Fructus and colleagues described a number of symptoms occurring during simulated dives between 300 and 365 msw. The symptoms have been termed the high-pressure nervous syndrome (HPNS) and include tremor in the hands and arms, increased slow-wave activity (2–7 Hz) and depression of α waves (8–13 Hz) in the EEG, dizziness, nausea, and vomiting. The syndrome was replicated by Bennett and Towse during a dive to 457 msw, and it has been demonstrated that lapses of consciousness may occur at depths greater than 300 msw. The symptoms of HPNS become more severe with increasing depth and during fast rates of compression.

Various methods have been used to ameliorate the signs and symptoms of HPNS, including the selection of less sensitive divers, the use of excursions from a shallow saturation exposure, slow exponential compression with stops for adaptation, and the use of narcotics to antagonize the effects of HPNS. Studies of these methods have shown that the speed of compression correlates with the extent and intensity of the various symptoms. In particular, a slow compression interrupted by holding stages has been shown to reduce HPNS considerably.

A major problem in evaluating the effects of HPNS by manipulating the compression conditions or selecting less sensitive divers is the individual variation in the reaction pattern. Whereas one diver may show a normal EEG but significant tremor followed by impaired motor performance, another diver may show normal motor functions but significant EEG changes accompanied by impaired cognitive performance. To study these different responses, it is recommended that a relatively comprehensive battery of tests be used for evaluating a compression profile. The validity of such a battery of tests can be related not only to fitness at depth but also to the temporal aftereffects that have been observed in deep diving.

Therefore, a compression profile with holding stages and a slower compression rate at deeper depth has been developed for operational diving to 180 msw and deeper in the North Sea. In a series of simulated dives to 360 msw, including a total of 18 divers, the proposed profile was evaluated. HPNS testing revealed only mild effects of the compression. Only three divers had impairments of more than 2 standard deviations in peripheral motor function. Memory was impaired periodically in two divers. Their tremor had minimal effects on their performance. In only l of the 18 divers had a marked EEG change followed by cognitive impairment. The results confirmed that using a compression profile with rates decreasing progressively with increasing depth, and with several intermediate stops, provides fit divers at depth.

This profile was later used in the tie-in operation at 220 msw at, by then, the world's largest oil platform (Gullfaks C) in 1989. A total of 52 divers took part in

the operation and all were tested with a field battery of HPNS tests at pressure before being allowed to start work in the water. No dives were aborted due to HPNS, and only 1 out of the 52 divers had to have an extended observation period in the living chamber (6 h) before being allowed to start the in-water operation. Although made even more conservative through two revisions, this original profile is the standard profile today for all diving in the Norwegian sector of the North Sea deeper than 180 msw.

Several studies have suggested that cerebral monoamines may be involved in the occurrence of some HPNS symptoms. Some studies suggested that dopamine (DA) counters the convulsive activity induced by pressure. Nevertheless, others did not find significant changes in the concentration or kinetics of DA analyzed in the brain homogenates of animals after exposure to pressure.

More recent studies in free-moving rats using microdialysis confirmed that the pressure induced striatal increase and implicated D1, D2, and NMDA receptors in this change. Recent works performed on $GABA_A$ or $GABA_B$ receptors using agonists (muscimol or baclofen) injected focally in the substantia nigra pars reticulata (SNr) or substantia nigra pars compacta (SNc) implicated these receptors in the striatal DA release and showed that helium pressure inhibits the activity of GABA neurotransmission at the level of substantia nigra, which induces a DA change.

Microdialysis studies indicated that the helium pressure induces an increase of serotonin (5-HT) in the striatum, which can explain the activation of these receptors. At the level of nucleus accumbens, the use of 5-HT-2C receptor agonists indicates that this receptor plays a crucial role in the development of locomotor activity (LMA) and in the pressure-induced increase in DA accumbens release. In contrast, the administration of 5-HT-2A receptor antagonists had no effect on LMA and DA increase, but it reduced myoclonia. These results suggest that helium pressure may simultaneously induce an increase in 5-HT transmission at the level of 5-HT-2A receptors and a decrease in 5-HT transmission at the level of 5-HT-2C receptors.

These studies are some examples of the several neuropharmacological and neurochemical findings that have increased our information on the origin and mechanisms of HPNS. In summary, it can be stated that helium pressure acts specifically on some neurotransmissions, on some neurochemical interactions between the systems and the structures, and on some receptors or on their synthesis. It is evident, however, that more work remains to be done to increase our understanding of the effect of pressure on the nervous system.

Isolation and confinement In addition to these pressure- and gas-related effects of deep diving on the CNS, studies have found that isolation and confinement themselves can have an effect. Although the stimulus-deprivation studies performed in the 1950s showed that the CNS needs stimulus variation, these extreme experiments seemed to be a long way from living in a small chamber (be it a saturation diving complex or a space station) for several weeks.

How has it been possible to evaluate this effect? Did someone pay to lock up test subjects inside a pressure chamber at normal atmospheric pressure for weeks? American, Russian, and European space agencies have used diving saturation chambers to evaluate the effects of long-term isolation and confinement in conjunction with preparing astronauts for space station Mir and the future international space station Alpha. Using these space-related experiments, it is possible to analyze the effects of isolation that a diver is exposed to during deep, long-duration (>30 days) dives. Three such experiments performed by the European Space Agency (ESA) are described next.

The Long-Term Programme Office (LTPO) of the Directorate of Space Station and Microgravity (D/SSM) at ESA carried out studies on the psychological and physiological aspects of long-duration human spaceflight in multinational crews under conditions of isolation and confinement in diving saturation chambers. LTPO first performed an experimental test with a crew of six men for 28 days (ISEMSI-90) at the Norwegian Underwater Technology Center (NUTEC) in Bergen, Norway, and then performed a second one with a crew of four (three men, one woman) for 60 days (EXEMSI-92) at Deutsche Zentrum für Luft-und Raumfahrt (DLR) in Cologne, Germany. In 1994, ESA was planning the EUROMIR-95 long-duration flight (135 days) of a European astronaut on board the Russian orbital station Mir. In order to perform detailed investigations on the human-related aspects in such a long-duration isolation, ESA organized a ground simulation, the Human Behavior Study (HUBES), for which the chamber facility at the Institute for Biomedical Problems (IBMP) in Moscow was set up like a small space station.

Cognitive fatigue and subjective health symptoms occurred in all three studies. On performance, an increase in decision time and an increase in the number of errors were found in several of the experiments. For the physiological studies, HUBES confirmed the earlier findings of the ISEMSI-90 and EXEMSI-92 tests, showing weekly rhythms, individual rhythms, and overall phases as in the group psychology studies. It was therefore concluded that for long-duration

dives (and spaceflights) experimenters will have to take into account the individual biological patterns of the subjects. If this is done, the outcome of the experiments will be more predictable and the comparison of data between the subjects will become easier.

These three ESA experiments have therefore shown, among other things, that the timing of individual experiments is critical when parameters are involved that are strongly under endocrinological control, such as those of the salt-water homeostasis, and therefore display rhythmic patterns.

Residual effects As in high-altitude medicine, researchers in diving medicine have vigorously debated the possibility of residual impairment of CNS function following the subjects' return to normal pressure. Several deep saturation dives performed between 1968 and 1989 showed temporary CNS changes immediately after the dives. Given these CNS changes, observed both immediately postdive and during deep saturation dives, an important question was whether chronic CNS effects could develop after single or repeated exposures, as in the previous findings on diving accidents.

The neuropsychological status of saturation divers was assessed before and after 300–500 msw dives (deep saturation diving, DSD group, $N = 64$) and before and after 3.5 years of ordinary saturation diving (SD group, $N = 32$). The average baseline results showed divers to be slightly superior to the nondiving controls. Mild-to-moderate neuropsychological changes (>10% impairment) were found in measures of tremor, spatial memory, vigilance, and autonomic reactivity in 20% of the divers after deep dives (DSD group). One year postdive, no recovery was observed except for a vigilance test. In the SD group, 20% of the divers showed >10% impairment after 3.5 years of ordinary saturation diving. A significant reduction in autonomic reactivity was also found, and there was a relationship between low autonomic reactivity before saturation diving and the number of impairments that were >10%. For the whole group (DSD + SD divers), negative correlations were found between saturation experience and results on memory and complex visuomotor tests. The number of years of diving from the first to last examination correlated positively with the number of impairments that were >10% and with a reduction in autonomic reactivity. No similar correlations were found at dive variables after approximately 3 years of air diving.

The mild-to-moderate changes seen in some divers, therefore, seem to be the effects of saturation diving. Because one deep dive may cause an effect similar to the effects of 3.5 years of ordinary saturation diving, there is reason to believe that repeated deep diving may lead to more pronounced neuropsychological impairment.

Such mild-to-moderate changes in some divers indicate the start of a pathological process on the subcortical level (mild encephalopathy) such as a decrease in memory and vigilance, low autonomic reactivity, and increased tremor, but the underlying mechanisms are unclear. The degree of individual reaction and the variation in symptomatology indicate that the mechanisms for both the temporal and the subtle permanent changes after saturation diving are very complex. Individual factors, additive factors, and synergistic mechanisms may be involved.

Several international studies followed the publication of the Norwegian study and were reviewed by Dutka. In his review, Dutka states that divers who have had decompression sickness can develop long-term health effects, although there are conflicting results among deep-sea divers with no history of decompression sickness. His conclusion is in conflict with the follow-up study on North Sea divers presented in 2003 by Troland. In this study, the results seem even more convincing, supporting the statement from an international consensus conference that said: "There is evidence that changes in bone, the central nervous system and the lung can be demonstrated in some divers who have not experienced a diving accident or other established environmental hazards. The changes are in most cases minor and do not influence the divers quality of health. However, the changes are of a nature that may influence the diver's future health (Hope et al. 1994)."

Humans Working under Low Pressure

Conditions

Human responses to hypoxia are complex and depend, among other things, on the severity of the hypoxia and the rate at which it is imposed. Severe sudden hypoxia experienced by aviators when cabin pressure is lost at a high altitude produces an entirely different result than that experienced by a climber who reaches the same altitude after weeks of acclimatization. Furthermore, people with a lifelong exposure to high altitudes develop further changes in response to hypoxia compared with even acclimatized lowlanders. Adaptations may also occur after a species has lived at high altitudes for generations due to the selection of genes advantageous to life in the hypoxic environment.

Generally, hypoxic conditions can be divided into four categories: (1) acute hypoxia (from seconds to

1–2 h), (2) chronic hypoxia (from hours to years, i.e., subjects in chamber experiments, mountaineers, workers in high-altitude mining towns, sea-level residents who go to high altitudes), (3) lifelong hypoxia (i.e., high-altitude residents who were born, bred, and live at high altitudes), and (4) species that have lived at altitude for generations.

Effects on the Central Nervous System

People who go to high altitudes often have changes in neuropsychological function, including memory, mood change, and special senses such as vision. This has been observed both in lowlanders staying at high altitudes for some time and in high-altitude natives. There is also mounting evidence that there may be persistent CNS defects on return to sea level after periods of severe hypoxia at high altitude. This is analogous to the indications of permanent CNS after-effects in some divers after deep diving.

Because an increasing number of climbers choose to climb at great altitudes without supplementary oxygen, these findings are of special interest. The increase in morbidity and mortality during such expeditions and the occurrence of irrational decisions made by severely hypoxic climbers have also led to serious concern from medical and operational points of view.

Acute hypoxia At modest altitude (>1500 m), breathlessness may be felt on exertion and some rise in heart rate may be noticed, but the main effect is on the CNS. Night vision is impaired at altitudes as low as 1500 m. The tingling of the fingers and mouth may be noticed at 4000–5000 m. However, although now definitely hypoxic, climbers have very little subjective prewarning. Some people may become unconscious above 5000 m, whereas most become unconscious at above 7000 m. In response to an acute exposure to Mt. Everest's altitude (8848 m), unacclimatized subjects remain conscious for only approximately 2 min.

Within the brain, the hippocampus, white matter, superior colliculus, and lateral geniculates appear particularly sensitive to levels of oxygen. Brain lactate levels increase during the early stages. The brain tissue concentrations of ATP, ADP, and AMP apparently remain close to normal even during severe hypoxia. In animal studies in which the arterial partial pressure of oxygen is reduced gradually (from 80 to 20 mmHg), the subjects' EEG amplitude initially increases slightly and then slow waves and spikes appear. Subsequently, the slow waves decrease in amplitude and then disappear; the small spikes become sporadic, and finally the EEG flattens. This initial activation followed by depression may be due to hypoxic effects on the reticular activating system. Visually evoked potentials are increased initially and are then abolished as the level of oxygen is reduced. The brainstem auditory response is abolished by hypoxia.

The effects of hypoxemia and hypocapnia at high altitude have opposing effects on the cerebral circulation. There have not, however, been systematic studies of the cerebral blood flow at various altitudes, partly because of the difficulties of measurement.

Chronic hypoxia at moderate altitudes It has been important to study the effects of mild hypoxia on the performance of aircrews, who are required to maintain peak performance for hours while performing monotonous tasks under hypobaric conditions. There is a general agreement that CNS function is impaired at altitudes over approximately 4500 m; however, what is the lowest altitude at which minor, but operational, significant alterations in CNS occur? This question arises frequently in the aviation industry because it is important in selecting the cabin pressure of commercial aircraft.

Earlier studies showed that performance declines as the time devoted to vigilance tasks increases. Studies on the effects of hypoxia on vigilance performance have produced contradictory evidence. Impairment in performance was found at altitudes as low as 4000 m and carboxyhemoglobin (HbCO) levels of l.8%. Other studies found no similar effects at 3500 m and at HbCO as high as 8.4%.

Further studies of people performing tasks at high terrestrial elevations have, however, demonstrated that impairment of a wide range of tasks occurs and that this correlates with the degree of hypoxia. In addition, animal performance in tests of maze performance, visual discrimination, or nondiscriminated avoidance are affected adversely by exposure to low oxygen tension. These changes are qualitatively similar to the behavior noted in rats when levels of brain biogenic amines, norepinephrine and DA, are reduced by pharmacological agents. Hypoxia also affects brain norepinephrine in rats. Exposure to 5 or 10% O_2 for l h resulted in a 23% decrease of this amine.

The variability in the response to hypobaric exposures, both within and between individuals, has been emphasized in relation to the symptoms and signs of hypoxia. Some proportion of the variation may be due to the differences in the respiratory responses to the hypoxia, causing significant temporal and individual differences in the tensions of oxygen and carbon dioxide in the arterial blood. Furthermore, the degree of cerebral vasoconstriction also causes a proportion of this variation.

Performing learned psychomotor tasks is unimpaired to at least 3000 m. Long- and short-term

memory, arithmetic, and conceptual reasoning tests are also unaffected below 3000 m. Previously, it was assumed that this degree of hypoxia did not produce any effects on human performance. However, breathing air at 2500 m causes a significant increase in the time required to achieve optimum performance on a novel task. The delayed learning is probably due to the retardation of some oxidative processes within the brain rather than to interference with oxidative phosphorylation, which hardly can be affected because the oxygen tension of the cerebral venous blood falls 2–4 mmHg during an ascent from ground level to 2500 m. The impairment of learning increases markedly at altitudes above 3704 m (12 000 feet).

A study by Værnes and colleagues was designed to examine the effects of a prolonged exposure to mild hypoxia on performance and endocrine reactivity. An altitude of 3048 m (10 000 feet) was simulated in an altitude chamber for 6.5 h. The study was initiated after reports of some cases of headache, dizziness, visual disturbances, and subjective feeling of performance decrement by an aircrew on a maritime surveillance aircraft during a trans-Atlantic flight with a cabin altitude of 3048 m. The aim of the study was to determine whether these symptoms were the effects of general stress, mild hypoxia, or a combination of both.

Performance tests indicated significant effects of hypoxia. In contrast to earlier studies on the relationship of hypoxia and performance, no relationship between impaired performance and duration of exposure to hypoxia was found. Repeated testing throughout exposure indicated stable individual reactions. The endocrine variables did not support the hypothesis that activation or stress caused the impairment observed. In addition to impaired neuropsychological test performance and impaired task performance, subjects reported headache, weakness, and some dizziness. Comparisons of the various tests confirmed the previous results showing that mild hypoxia yields varying degrees of impairment on different cognitive functions.

Changes that take place over a period of hours to months at altitude, termed acclimatization, occur in different systems of the body and have a different time course. In some cases, they involve a biphasic response (i.e., the heart rate shows a rise within minutes, followed by a fall over weeks at altitude). The best-known response is the increase in hemoglobin concentration due to the rapid decrease in plasma volume followed by a much slower increase in red cell mass.

Adaptation takes place over decades or is the result of being born and bred in a mountain environment. No definite examples of CNS-related adaptation have, however, been demonstrated in humans, but examples include an increase in the number of branches leaving the main coronary trunks and the blunting of the hypoxic ventilatory response.

The observed deterioration in the mental and physical conditions of humans as a result of prolonged stay at altitude is termed high-altitude deterioration. The term was first used by members of the early Everest expeditions. High-altitude deterioration, in the absence of dehydration, hypothermia, exhaustion, and/or starvation, is characterized by poor appetite, weight loss, lethargy, slow recovery from fatigue, an increasing lack of will power to start new tasks, and irritability. The mechanisms underlying this deterioration are unknown, but it is quite obvious that the effect of chronic severe hypoxia on the CNS plays an important part.

Chronic hypoxia at high altitudes Both simple and more complex psychological functions are impaired significantly at high altitudes (>6000 m), including arithmetical tests, writing ability, and the appearance and disappearance of afterimages following exposure of the eye to bright light. Increased memory errors, errors in perseverance, and reductions in auditory threshold and words apprehended are also observed. In a series of now-classic studies, measurements of sensory and motor functions were made at several altitudes. It was found that impairment was not significant below 5300 m. Subjects with longer periods of acclimatization appeared to suffer less deterioration.

The general findings from several such performance studies showed consistently that there is an increased distractibility and lethargy that tend to reduce the subjects' ability to concentrate. These results are consistent with the impression reported by many people who have worked at high altitude, namely that thought processes are slowed but that, if one concentrates hard enough, accurate procedures can be carried out.

Changes in the human EEG have been observed at high altitudes (5500 m). Ryn reported both an α-wave. frequency decrease and paroxysmal and focal pathology. In general, there were abnormalities in 11 out of 30 climbers. Partly the same pattern was later found by Zhongyuan and colleagues at altitudes above 5000 m. However, the changes were less than those observed during acute hypoxia at the same altitude simulated in a pressure chamber prior to the expedition. A general finding was that members of the expedition who tolerated the acute hypoxia well tended to show fewer EEG changes on the mountain.

Recent studies on altitude exposure above 6000 m showed an increased susceptibility to hallucinatory

experiences in mountaineers who refused to use supplementary oxygen. The autonomic nervous system seems to play a crucial role in the short-term adjustment to hypoxia. There is increased sympatho-adrenal activity associated with decreased parasympathetic tone, and vasomotor depression due to central hypoxia may occur in nonacclimatized subjects with reduced altitude tolerance at and above 4000 m.

Findings of possible impairment of higher cognitive functions at high altitude are still controversial. Recent studies have more specifically studied what is sometimes called mountaineer's high, which in some ways resembles the syndrome associated with lesions in the frontolimbic loops. The similarity in the patterns of cognitive and emotional deficits after frontal lobe impairment and in the high-altitude syndrome is best accounted for by the fact that the prefrontal and orbitofrontal areas and their connections to the limbic centers are among those most sensitive to acute cerebral hypoxia. When three groups of seven men were exposed to simulated altitudes of 3000 and 4500 m, short-term adaptation mechanisms led to a preservation of selected frontal lobe cognitive and emotional functions. Functional hemispheric superiorities for word recognition and affective processing were likewise unaffected, despite a drop in diastolic blood pressure at 4500 m, indicating a beginning central hypoxia in terms of functional impairment of the vasomotor center.

Residual effects As in diving medicine, those in high-altitude medicine express great interest in the possibility of residual impairment of CNS function following a return to normal pressure. Studies were done of 21 members of an American expedition to Mt. Everest in 1981 prior to the expedition, immediately afterward, and 1 year later using a comprehensive battery of neuropsychological tests. Changes were found for verbal learning and memory, aphasic errors, and fine-motor tempo. Other authors reported similar or consistent findings. Results from a second American expedition in 1985 confirmed the changes found in the 1981 expedition. Furthermore, a simulated altitude exposure similar to the Mt. Everest conditions produced the same impairments in motor speed and persistence, memory, and verbal expression afterward. There was also a significant negative correlation between hypoxic ventilatory response and neuropsychological function measured after the expedition. In other words, subjects with the largest hypoxic ventilatory response showed the greatest neuropsychological decrement. Townes explained this finding by the following hypothesis. Subjects with the highest hypoxic ventilatory response would reduce their arterial pressure of CO_2 the most and therefore develop the most cerebral vasoconstriction. As a result, this would cause the most severe cerebral hypoxia, even though their arterial pressure of O_2 would actually be higher than in subjects with the smaller ventilatory response to hypoxia. Ward and colleagues found that this correlation between hypoxic ventilatory response and residual impairment of CNS function led to a paradox: the climber who is endowed by nature to go the highest is likely to suffer the most severe CNS damage.

Concluding Remarks

Both diving and high-altitude climbing are potentially dangerous. It requires time and patience to master all the skills necessary to move safely at deep depths and in mountain country. Vigilance and emotional stability are important, and extreme sensation seekers are not as likely to function effectively as more self-sufficient people. Those who are obliged to live harmoniously in close proximity for long periods of time in a high- or low-pressure environment should be stable, loyal, and have both a social and a intellectual tolerance for their companions.

Further Reading

Behnke, A. R. (1942). Effects of high pressure prevention and treatment of compressed air illness. *Medical Clinics of North America* **26**, 1228–1237.

Behnke, A. R., Thomson, R. M. and Motley, E. P. (1935). The psychologic effect from breathing air at 4 atmospheres pressure. *American Journal of Physiology* **161,** 417–425.

Bennett, P. B. (1982). The high pressure syndrome in man. In: Bennett, P. B. & Elliott, D. H. (eds.) *The physiology and medicine of diving*. London: Baillière Tindall, 262–296.

Brauer, R. W., Dimov, S., Fructus, X., et al. (1969). Syndrome neurologique et electrographique des hautes pressions. *Revue Neurologique* **121**, 264–265.

Brugger, P., Regard, M., Landis, T., et al. (1999). Hallucinatory experiences in extreme-altitude climbers. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology* **12**, 67–71.

Burnett, C. S. F. (1996). High altitude mountaineering 1600 years ago. *Alpine Journal* **88**, 127.

Cavaletti, G., Morom, R., Garavaglia, P., et al. (1987). Brain damage after high-altitude climbs without oxygen. *Lancet* **1**, 101.

Clark, J. M. (1982). Oxygen toxicity. In: Bennett, P. B. & Elliott, D. H. (eds.) *The physiology and medicine of diving*. London: Baillière Tindall, 200–238.

Cousteau, J. Y. (1953). *The silent world*. London: Reprint Society.

Darbin, O., Risso, J. J. and Rostain, J. C. (1999). The full expression of loco motor and motor hyperactivities

induced by pressure requires both striatal dopaminergic and NMDA receptors activities in rat. *Neuroscience Letters* **267**, 149–152.
Donald, K. W. (1947). Oxygen poisoning in man. *British Medical Journal* **1**, 667–672.
Dutka, A. J. (2003). Long term effects on the central nervous system. In: Brubakk, A. O. & Neuman, T. S. (eds.) *Bennett & Elliott's physiology and medicine of diving*. Edinburgh: Saunders, 680–699.
Hope, A., Lund, T., Elliott, D. H., Halsey, M. J. and Wiig, H. (eds) (1994). *Longterm health effects of diving: an international consensus conference*. Bergen: Norwegian Underwater Technology Center, 391.
Værnes, R. J., Hammerborg, D., Ellertsen, B., et al. (1983). Central nervous system reactions during heliox and trimix dives to 51 ATA, Deep Ex 81. *Undersea Biomedical Research* **10**, 169–192.
Værnes, R. J., Kløve, H. and Ellertsen, B. (1989). Neuropsychologic effects of saturation diving. *Undersea Biomedical Research* **16**, 233–251.
Værnes, R. J. and Sandal, G. (2003). Human reactions to deep-water conditions. *Lancet* **362**, 10–12.
Ward, M. P., Milledge, J. S. and West, J. B. (1989). *High altitude medicine and physiology*. London: Chapman and Hall.

Primate Hierarchies and Personality

R M Sapolsky
Stanford University, Stanford, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R M Sapolsky, volume 3, pp 221–225, © 2000, Elsevier Inc.

Background: Primate Sociality
Social Rank and Stress Physiology
Modifiers of the Rank–Physiology Relationship
Hierarchies, Humans, and Stress-Related Disease

Glossary

Primate personality	Stable patterns of behavioral style, emotional responsiveness and social roles.
Dominance hierarchies	The collectivity of an array of reasonably stable assymetries in dyads of animals is competing for limited resources.
Psychosocial stressors	Stressors, in a social context, in which there is either anticipation of a challenge to homeostasis (without it having yet occurred) or in which there is no threat of a challenge whatsoever.

Whether one is a professional zoologist or a casual observer of animals in a zoo or on a public television special, few can resist the pull of non-humans primates. Part of the fascination with them is due to their intelligence, and part to their unnervingly close relationship to us. Perhaps the strongest pull, especially when considering the apes and the Old World monkeys, is the force of their individuality – these are long-lived organisms with strong, pungent personalities and temperaments. For the psychobiologist, an irresistible question is whether there are physiological correlates of these individual behavioral differences and, if such physiological differences exist, whether they are causes or consequences of the behavior.

Stress physiologists have examined this question for a long time, framing it in the context of the stress response and stress-related disease. Do individual differences in primate behaviors, in social rank within the dominance hierarchy, or in personality predict differences in the quality of the stress response? Do they predict differences in patterns of stress-related disease? The answers to these questions are complex, and the very complexity is what offers the greatest insight into the human condition. To best appreciate the complexity, a brief overview of primate social systems will be helpful.

Background: Primate Sociality

As befits our close phylogenetic relatives, the primate order is arguably the most flexible, varied, and catholic of any in the animal kingdom. Primates can be found in ecosystems ranging from the most severe of deserts in Africa to snowy regions of north India and Japan, from thick rainforests to open savanna, from flat plains to steep mountainous regions. They are equally varied in their diet, ranging from strict vegetarians to organized hunters. Given this heterogeneity, it is not surprising that there is an enormous range of primate social systems, ranging from solitary, to small related groups built around male–female pair bonds, to fusion/fission societies involving hundreds of individuals with complex patterns of

polygamy and polyandry. Nonetheless, there are some themes common to most primate species.

First, with the exception of the few rare solitary primate species (such as the orangutan), these are profoundly social animals; the behavior of most primates outside the context of their group makes little sense. As but one measure of the importance of sociality, across the more than 150 primate species, the size of the cortex (relative to total brain weight) positively correlates with the size of social groups in that species. Second, most primate societies have exogamy of one sex, that is, the adolescents of one sex leave the group at puberty, transferring to another group as an inbreeding-avoidance mechanism. Among chimpanzees, for example, it is females who transfer, whereas it is males among the baboons and macaques. As a result, the gender that does not transfer spends its entire life in its natal group, surrounded by same-sex relatives and its mother. Thus, within any such primate society, it is impossible to understand the behavior of one of the sexes without considering the complexities of interactions among kin – cooperative coalitions among chimpanzee brothers or baboon sisters, aunting behavior by female macaques that increases the likelihood of a sister's child surviving, and so on. Finally, even in the most bountiful of ecosystems, resources are not infinite and, in most primate societies, are typically divided up unevenly along lines of conventionalized asymmetry referred to as dominance hierarchies. In a stable dominance hierarchy, limited resources are not contested with bloody tooth and claw. Instead, animals are smart enough to understand the status quo of the dominance hierarchy, allowing it to substitute for overt competition. Accordingly, the social rank of an animal affects its quality of life to an enormous extent.

The phenomenon of dominance hierarchies has long fascinated primatologists, as it can be the most immediate and dramatic version of individual differences in the behavior of these animals. Thus, when stress physiologists became interested in the psychobiology of individual differences among primates, they first focused on social rank.

Social Rank and Stress Physiology

The earliest theorizing in the field was badly derailed by the influence of the executive stress syndrome studies of the 1950s. This work seemed to suggest that monkeys with executive control over resources were more prone toward ulcers than monkeys without control, and was widely interpreted as evidence of the increased stressfulness of dominance. These findings were subsequently shown to be spurious, due to nonrandom selection of subjects. Specifically, in order to facilitate the learning of the operancy performed as the executive task, animals were prescreened, and the most emotionally reactive ones were made executive animals.

Finally, freed from these erroneous findings, primatologists interested in the rank–health relationship turned their attention to rodent studies, which were pioneered by James Henry. These produced a fairly consistent picture. Subordinate animals (as typically detected by the outcome of paired fights) showed indices of chronic stress, including elevated circulating glucocorticoid levels, hyperplastic adrenals, immune suppression, reproductive problems, and hypertension. As studies began with non-human primates, it was expected that the same would be seen. This was reinforced by insights into the psychological variables that modulate the stress response. Specifically, classic studies by John Mason, Jay Weiss, Seymour Levine, Martin Seligman, and colleagues showed that the same physical stressor would trigger a larger or more prolonged stress response if the organism lacked a sense of control, lacked a sense of outlets, lacked predictive information about the stressor, or interpreted events as worsening.

Studies of an array of primate species demonstrated that lower-ranking individuals in stable dominance hierarchies disproportionately suffer those psychological stressors. They lack control and predictability, for example, potentially spending minutes obtaining a food item that can be seized from them, finding a grooming partner only to have the bout disrupted by someone else, being subject to unpredictable displacement aggression if someone else was having a bad day. Moreover, they typically lack the means to release frustration by displacing aggression onto someone else.

This generated the expectation that subordinate primates would have a maladaptive profile that would include stress responses that were hyperactive basally or sluggish in recovering from a stressor, and greater vulnerability to stress-related disease. This was precisely what was reported, during the 1980s and 1990s. Work by, among others, Jay Kaplan, Carole Shively, Deborah Gust and colleagues with macaques, Robert Sapolsky with baboons (a wild population, in this case), Barry Keverne and colleagues with talapoins, Kirk Manogue and colleagues with squirrel monkeys, and Eberhart Fuchs and colleagues with tree shrews demonstrated that subordinate individuals displayed elevated basal glucocorticoid levels and negative feedback insensitivity, elevated resting blood pressure, a sluggish cardiovascular response to a stressor and a sluggish recovery, suppressed levels of good (i.e., HDL) cholesterol, impaired fertility and a

reproductive axis more readily suppressed by stress, fewer circulating white blood cells, an enhanced risk of atherosclerosis, increased endogenous benzodiazepine signaling, and less genesis of new neurons in the brain. Furthermore, when group membership was being manipulated among captive animals, and physiological measures were taken before and after formation of the group (and thus of the dominance hierarchy), it was clear that the physiological differences arose as a consequence of the rank differences. Finally, an array of subtle studies showed that these stress-related indices were far more a function of the psychosocial aspects of subordination, rather than one of the overt physical insults of rank.

These findings were readily interpreted as indices of the stressfulness of social subordination. A number of individuals carried out detailed studies uncovering the reductive mechanisms explaining these patterns. For example, Kaplan and colleagues have done detailed studies demonstrating that the stress-induced suppression of estrogen levels in subordinate female macaques helps contribute to the increased risk of atherosclerosis. Sapolsky and colleagues have shown that basal hypercortisolism of the subordinate baboons arises for many of the same reasons that hypercortisolism does in human depression (i.e., hypersecretion of corticotropin-releasing hormone [CRH] at the hypothalamic level, a partial compensation at the level of the pituitary via blunted sensitivity to CRH, and feedback resistance in the form of insensitivity to dexamethasone).

Despite this seeming clarity, major contradictions to this picture were emerging. David Abbott, working with marmoset and tamarin monkeys, and Cavigelli, working with feral lemurs (as well as Scott Creel, studying an array of non-primate species, including African wild dogs and dwarf mongooses), were reporting diametrically opposite findings, namely, that it is dominant individuals that show the indices of overactivated stress responses.

By the mid-1990s, a widespread impression was emerging that there was, in fact, little consistent relationship between primate social status and stress-related disease. The most recent work in the field has shown that this is not the case. Instead, the rank–health relationship can be modified dramatically by an array of additional factors.

Modifiers of the Rank–Physiology Relationship

What Is the Meaning of Rank in a Species?

Among the macaques and baboons, subordination is a state forcibly imposed from above by way of aggression or harassment by dominant individuals; resources are divided with marked inequity along the lines of rank. Such despotic dominance systems generate considerate stress for subordinates. The situation for marmosets and tamarins, in contrast, is quite different. Among these New World monkeys, social groups of six to eight animals consist of a pair-bonded dominant pair and a number of subordinate individuals. Critically, these social groups are typically extended families of relatives, and successful reproduction involves cooperativity among them. A subordinate animal is usually a younger relative waiting her turn and helping older relatives with childcare in the interim, and is rarely subject to displacement aggression. This goes far to explain Abbott's reports of dominant animals showing the preponderance of stress-related physiology. One meta-analysis examined all studies of rank–basal glucocorticoid relationships in non-human primates (seven species, both genders), ranging from cases in which subordinate animals have far lower basal levels than dominant animals (marmosets and tamarins) to those in which subordinates have far higher (talapoins and baboons). Across these species, a significant predictor of such relative glucocorticoid levels was the degree to which subordinates were subject to the harassment and displacement aggression typical of despotic dominance systems.

How Is High Rank Maintained in the Particular Primate System?

In many primate social systems, rank can change over time, raising the question of how a high-ranking individual holds on to that rank as long as possible. In some cases, this involves frequent and overt aggression aimed at subordinate individuals, with the result that dominant animals are involved in the highest rates of fighting. In contrast, in other cases, well-entrenched dominant animals have the lowest rates of aggression, maintaining their dominance through little more than subtle psychological intimidation. A number of studies now demonstrate that in the former case, subordination is not associated with elevated basal glucocorticoid levels, whereas in the latter it is.

How Stable Is the Dominance Hierarchy?

In many primate social systems, rank is inherited at birth, remaining static throughout life. However, even in systems in which rank shifts over time, there can be long periods of hierarchical stability (i.e., where the dominant individual in any given dyad wins the vast majority of dominance interactions, thereby reinforcing the status quo). In such cases,

dominant individuals have the abundant physical and psychological advantages of their lofty rank. Occasionally, hierarchies will become unstable (due to some key individual entering or leaving the troop, a death or injury, or the formation of dissolution of a key coalition), and there will ensue an unstable period of dramatic, frequent shifts in rank until the hierarchy restabilizes. This period will produce high rates of aggression and competitive interactions among individuals, unpredictable formation and disintegration of coalitions, and low rates of socially affiliative behaviors. Under such circumstances, it is high-ranking individuals, at the center of the competition, who have the least control and predictability over events around them. As shown by numerous investigators, during such unstable periods, it is no longer the subordinate individuals who have the highest basal glucocorticoid levels or the greatest risk of developing atherosclerosis.

What Coping Strategies Are Available to Subordinate Animals?

The magnitude of the stress response is a function of the severity and frequency of stressors to which an individual is exposed and, just as importantly, the availability of coping outlets. In primates, some coping outlets are positive (grooming, being groomed, physical contact with another animal) and others are negative (displacing aggression on to another animal). Across the primate species studied, the more such outlets are available (both positive and negative), the lower the basal glucocorticoid levels. There are not sufficient data, however, to determine whether positive or negative outlets are more physiologically protective.

How Readily Can Subordinate Animals Avoid Dominant Ones?

In a rather simple and intuitive way, subordinate animals who can readily evade the attention of dominant individuals are likely to be subject to fewer stressors. This point helps explain some seeming contradictions in the literature. For example, studies of wolves have shown that in captivity (where subordinate individuals are markedly constrained in their ability to avoid dominant animals), it is subordinate animals who have the highest basal glucocorticoid levels. In contrast, in feral wolf populations, this is not the case.

What Is the Social Milieu of a Particular Primate Group?

Primate social systems not only vary between species, but also vary among different populations of the same species. In some cases, this reflects ecological or genetic issues, but often, it reflects a fluke of demographics of the group and the personalities of its members. As such, the social milieu of a particular group can modify the rank–health relationship. Work by Gust and by Sapolsky has focused on social systems (of female macaques and male baboons, respectively) in which subordination is typically associated with elevated basal glucocorticoid levels. However, this is not observed in troops with atypically low levels of displacement aggression on subordinates and/or high levels of social affiliation.

What Is an Individual Animal's Experience of Life in Its Social Group?

While the generic features of a primate's social group (e.g., rates of reconciliation, extent of hierarchical stability) may be a significant predictor of stress-related physiology, so is the personal experience of those variables. For example, in a pair of studies of macaques, the severity of basal hypercortisolism varied as a function of how often individual animals were subject to dominance or aggressive interactions, or how often they were given affiliative support. Similarly, in a study of wild baboons during a period when the hierarchy was destabilized by a highly aggressive transfer male, the frequency with which that male attacked females was highly predictive of the extent to which they were immunosuppressed. Strikingly, in that study, females who had not been subject to such displacement aggression showed no evidence of immune suppression. Thus, these stress-related physiological markers do not respond to the abstract features of life in a particular troop (i.e., this is a troop with high rates of reconciliation/displaced aggression/coalitional fighting ...), but to the individual's very concrete personal experience of events in the troop.

What Is the Personality of the Animal in Question?

One of the most striking factors influencing individual differences in stress-related physiology among primates is personality. The notion of personalities among non-human primates was long ignored in many circles of primatology, being viewed as unacceptably anecdotal and anthropomorphic. Recent and rigorous work has shown there to be consistent and stable differences among individuals in their temperament, patterns of impulsivity, tendencies toward sociality, the social roles they will assume within the group, and so on. With those findings has come an interest in whether those differences predict different profiles of stress-related physiology.

As one example, Stephen Suomi and colleagues, Ned Kalin, as well as Kaplan, Steven Manuck, and colleagues reported stable personality differences in macaque monkeys in their extent of reactivity to novelty. Such traits are apparent very early in life, and the high reactor animals whose behaviors are more disrupted by novelty are hypercortisolemic throughout life and more prone toward atherosclerosis.

In another series of studies, Sapolsky and colleagues examined personality differences among wild baboons, after controlling for rank. Two clusters of traits predicted elevated basal cortisol levels. The first related to patterns of male–male competition; hypercortisolism was observed among individuals who were least capable of differentiating between threatening and neutral interactions with rivals (i.e., they responded to both as if they were provocative), were least likely to control the initiation of a fight, and were least capable of differentiating between winning and losing fights (i.e., both outcomes equally disrupted their social behavior thereafter). A second, independent cluster demonstrated hypercortisolism among the males with the lowest rates of social grooming or contact with females, or play with infants. These findings were interpreted as showing greater endocrine indices of stress among animals with the least amounts of control and predictability about stressors and the fewest sources of social coping outlets.

Hierarchies, Humans, and Stress-Related Disease

A number of investigators have examined human hierarchies in the hopes of correlating rank differences with differences in stress physiology. Outcomes of athletic competitions, military ranks, and position within a corporate workplace have all been considered as examples of human hierarchies, and differences in physiology have been reported. Other investigators have criticized these studies due to artificiality of these rankings (e.g., how much can rankings in a college wrestling tournament tell us about cardiovascular disease susceptibility?), the confounding ability of humans to be in more than one hierarchy (i.e., the low-ranking individual working in the mailroom of a corporation might, at the same time, be captain of the company softball team and consider that to be the more relevant hierarchy), and the human capacity for rationalizing and externalizing what would be perceived as being a low rank.

Some investigators have suggested that low socioeconomic status (SES) can be viewed as the most valid and pervasive example of low rank in a human, carrying with it many of the psychological stressors seen among many low-ranking non-human primates (lack of control, predictability, or outlets). Low SES carries with it an enormously increased risk of a broad range of diseases, and careful study by pioneers such as Michael Marmot, Richard Wilkinson, Nancy Adler, and Ichiro Kawachi has shown that this SES gradient can not be explained by differences in health-care access and can only be minimally explained by differences in education or in exposure to risk factors and protective factors. In light of those findings, some researchers have suggested that part of the SES gradient of disease can be viewed as reflecting the psychosocial stressors of poverty. In support of this is not only the fact that some of the alternative explanations can be ruled out to some extent (e.g., a striking bit of evidence against the idea that the SES gradient is due solely to differential health-care access is the robust presence of such gradients in countries with universal health care), but also the fact that diseases that are probably the most stress sensitive (psychiatric disorders and cardiovascular disease) show the steepest gradients.

See Also the Following Articles

Health and Socioeconomic Status; Primate Models, Overview; Social Networks and Social Isolation; Social Status and Stress; Social Stress, Animal Models of.

Further Reading

Abbott, D., Keverne, E., Bercovith, F., et al. (2003). Are subordinates always stressed? A comparative analysis of rank differences in cortisol levels among primates. *Hormones and Behavior* **43**, 67.

Clark, A. and Boinski, S. (1995). Temperament in nonhuman primates. *American Journal of Primatology* **37**, 103.

Creel, S. (2001). Social dominance and stress hormones. *Trends in Ecology and Evolution* **16**, 491.

Gould, E. and Gross, C. (2002). Neurogenesis in adult mammals: some progress and problems. *Journal of Neuroscience* **22**, 619–623.

McEwen, B. (2002). Sex, stress and hippocampus: allostasis, allostatic load and the aging process. *Neurobiology of Aging* **23**, 921.

Sapolsky, R. (1999). Stress, glucocorticoids and their adverse neurological effects: relevance to aging. *Experimental Gerontology* **34**, 721.

Sapolsky, R. (2000). Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. *Archives of General Psychiatry* **57**, 925.

Sapolsky, R. (2005). The influence of social hierarchy on primate health. *Science* **308**, 648.

Suomi, S. (1997). Early determinants of behaviour: evidence from primate studies. *British Medical Bulletin* **53**, 170.

Primate Models, Behavioral–Immunological Interactions

M L Laudenslager and S Tiefenbacher
University of Colorado Health Sciences Center Denver, CO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M L Laudenslager, volume 3, pp 226–229, © 2000, Elsevier Inc.

Animal Models
The Immune System
Prenatal Experiences Modify Postnatal Immune Competence of the Offspring
Postnatal Experiences Have Acute and Long-Term Consequences
Social Relationships Affect Immunoregulation
Individual Differences in Immune Regulation
Environmental Factors Influencing Immunoregulation
Mechanisms of Brain–Immune Interactions
Immune Senescence in Nonhuman Primates
Summary

Glossary

Acquired immunity	The components of the immune system that require time (several days) to respond to antigens and are highly specific for an antigen.
Antibody	A protein secreted by B cells that binds to antigens and leads to their destruction and removal, also called immunoglobulin.
Antigen	Any substance that is recognized by the immune system as not belonging to the body (not self); includes viruses, bacteria, tumors, and tissue grafts.
Cytokines	Protein messengers made by a variety of cells including lymphocytes and phagocytes that regulate the immune response and affect central nervous system activity.
Innate immunity	The components of the immune system that respond immediately in the presence of antigens as a rapid nonspecific response, a first line of defense; includes neutrophils, natural killer T cells, and macrophages.
Lymphocytes	The cells of the immune system, including B cells (make antibodies), T cells, (regulate the immune response and destroy foreign cells), and natural killer (NK) cells (first line of defense against tumors and virally infected cells; also regulate the immune response).
Phagocytes	The cells of the immune system that process antigens in preparation for activating the immune response; include neutrophils and macrophages.
Psychoneuroimmunology	An interdisciplinary science that examines bidirectional interrelationships among immune, neuroendocrine, and behavioral processes
Simian immunodeficiency virus (SIV)	A virus found only in monkeys that is the best model for human immunodeficiency virus (HIV) available to scientists.

In this article, the utility of animal models is described with a focus on nonhuman primates, followed by a simple overview of the immune system. The area of immunology is growing by leaps and bounds, and as such it is elusive with regard to any review that serves the entire discipline with any justice. The growing area of psychoneuroimmunology (PNI) suggests that there are important regulatory relationships between the brain and the immune system and furthermore that these relationships are bidirectional. The study of these relationships in nonhuman primates is of particular importance for improved understanding brain immune relations in humans due to the great physiological and behavioral similarity of the two groups. The following discussion focuses on several areas of investigation that have implications for immune and behavior relationships in nonhuman primates: prenatal and postnatal experiences, social relationships, and aging.

Animal Models

Animal models are used in a variety of biomedical studies because they permit the control of a number of factors not possible with human subjects. An adequate animal model of any human condition must fulfill several criteria, including a common phenomenology, etiology, underlying pathophysiology, and effective treatment. The evolutionary proximity of humans and nonhuman primates (e.g., monkeys and apes) makes the nonhuman primate a frequently studied model for both human health and disease. The most obvious similarity to humans resides in the

social nature and behavioral characteristics of nonhuman primates. In addition, nonhuman primates offer a unique opportunity for studying brain–immune relationships. It is likely that the phenotypic traits shared by humans and nonhuman primates are homologous, that is, that they are related to a common ancestral origin. Thus, they may share similar design features. These similarities are important for disentangling complex biomedical problems.

The Immune System

The immune system is a highly regulated system with numerous redundancies and backups. Simply put, the responsibilities of the immune system can be characterized by the three Rs: recognition, removal, and regulation. The immune system must recognize or differentiate self from not self (antigens or foreign substances) and respond appropriately. The antigen might be a simple microorganism located in a cut to a tumor developing in the lungs. Once an antigen is recognized as not self, it must be removed or destroyed by host defense mechanisms. This very complicated process is regulated by immune cells that communicate via complex proteins called cytokines, endocrine hormones, and neurotransmitters derived from neural structure. There are innate processes that respond immediately to the antigen until the acquired system can respond to the foreign intruder. The ability of natural killer (NK) cells to immediately destroy tumors or virally infected cells illustrates the actions of the innate immune system. The presence of specific antibodies protecting the individual against flu, following the administration of a flu vaccine in the fall, exemplify the actions of the acquired immune response. Acquired immunity is a slower and highly specific process; flu protection after vaccination does not occur until approximately 14 or more days after a person receives the vaccine. This response is also a highly integrated response of the immune system.

These systems are not independent but act together closely to control infection and disease. The innate system forms a first line of defense until the acquired system can develop responses specific for the pathogen. It might take from several days to weeks before the acquired response is at full capacity. The innate and the acquired response are regulated by the actions of cells of the immune system, including lymphocytes (T cells, B cells, and NK cells) and phagocytes (neutrophils, macrophages, and eosinophils). T cells, NK cells, and macrophages secrete a number of protein messengers (cytokines) that are critical in the regulation of the response. The capabilities of the immune system can be measured in either tissue culture systems (*in vitro*) or in the intact organism (*in vivo*). *In vivo* measurements are preferred over *in vitro* assessment because the system is studied under naturally occurring conditions. For example, specific antibodies seen after immunization with a foreign protein represent the final output of a highly integrated response of the immune system. The antibodies in the blood reflect the initial recognition and processing of the antigen by phagocytic cells; release of regulatory cytokines by T cells, NK cells, and macrophages; and interactions with neural and hormonal factors resulting in the final production of antibody by the B cells. One of the first reports on stress in nonhuman primates described the suppressive effects of unpredictable and generally poor housing conditions of monkeys on their antibody levels after immunization.

Prenatal Experiences Modify Postnatal Immune Competence of the Offspring

The prenatal environment contributes significantly to postnatal neurobehavioral developmental and immune competence and as such is a significant contributor to individual differences in the overall regulation of the hypothalamic-pituitary-adrenal (HPA) axis and immune system. The mother and the developing fetus form a highly regulated amalgamation that has received considerable attention with regard to the fetal origins of later medical conditions in the offspring. Stress in the mother is synonymous with stress in the developing fetus. More specifically, prenatal stressors (behavioral and/or pharmacological, e.g., dexamethasone or adrenocorticotropic hormone, ACTH) of pregnant rhesus monkeys are associated with a cascade of postnatal effects in the fetus. The power of prenatal stressors is indicated by the fact that they are sufficiently pervasive to redirect the gender-specific development of the offspring. Although the offspring are healthy and normal at birth, the effects of prenatal stress are reflected in a dysregulation of immune responses and a more responsive HPA axis in infants born to these prenatally challenged mothers. Immune effects include suppressed cytotoxic-like activity of lymphocytes, depressed proliferative responses, and reduced cytokine production by lymphocytes. The timing of the stressor during gestation (e.g., early vs. late) affects the direction of the specific impact on immunoregulation for some of these measures. That is, early gestational stressors are associated with enhanced mixed lymphocyte responses (MLR) responses toward syngeneic cells from the mother or father, whereas late gestational stressors are associated with reduced MLR to the same targets. Thus, the nature of prenatal stressors is complicated but nonetheless powerful. Prenatal stressors impact placental transfer from

mother to fetus, affecting the transfer of maternal immunoglobulins and iron. These effects lead to reduced levels of beneficial bacteria in the gut of the offspring.

Postnatal Experiences Have Acute and Long-Term Consequences

Early experience effects are not limited to events during the prenatal period but extend to the postnatal period as well. Behavioral challenges are similar in their impacts on later immune and HPA regulation. For example, significant life challenges such as brief maternal separations that might occur during the breeding season when mothers leave their offspring to mate with males in the troop, matrilineal dominance status, or the quality of maternal care have significant effects on the immune regulation in nonhuman primates, both in the field and in laboratory settings. Wild-born offspring of low-ranking matrilines, for example, show greater immune and HPA perturbations to acute challenges as adolescents. In the laboratory, social-separation experiences such as the removal of the mother from the infant for brief periods of time produce a number of disturbances in both *in vitro* and *in vivo* immune parameters. Acutely, *in vitro* immune measures such as NK activity and activity of macrophages are elevated, whereas other markers are suppressed. Indeed, specific antibody levels following immunization at the time of separation are reduced and the magnitude of this reduction is greatest in the most distressed monkeys. These early separation experiences have also been shown to impact immune regulation over an extended time period, such as reduced *in vitro* markers of the T-cell function and an elevation in markers of the innate system, namely NK activity and the activity of the macrophage system.

Postnatal pharmacological challenges such as antidepressant drugs administered to young monkeys can have long-lasting effects on immune competence, resulting in lower specific antibody responses to antigenic challenge and increased reactivity of the HPA axis. Altered HPA regulation is one of many contributing factors to dysregulated immune function. The overarching observation is that, if something stressful happens to the mother–infant unit prenatally or the infant postnally, it is very likely to result in a disruption of the immune system, sometimes in subtle but important ways. These observations also indicate the utility of nonhuman primate models, with their shortened life cycle and the researchers' ability to control environmental exposure to pathogens that will surely impact the developmental trajectory. This makes the nonhuman primate model ideal for studies of aging.

A critical question facing the field of psychoneuroimmunology is the interpretation of the functional significance of these *in vitro* measures. As addressed earlier, *in vivo* measures are most likely to reveal important consequences mediated by immune activity that might not be indicated by *in vitro* measures. The implications of early challenges and stressor experiences are best exemplified by studies of monkeys infected with SIV, a nonhuman primate model for acquired immunodeficiency syndrome (AIDS). There is considerable variability in the disease course of AIDS after infection with human immunodeficiency virus (HIV). A significant question for understanding the pathophysiology of this virus is: Why do some individuals infected with HIV progress rapidly and become quite ill and others infected with the virus remain asymptomatic? In retrospective studies of rhesus monkeys infected with SIV, researchers found that the early developmental history of the monkeys contributed significantly to disease course. Thus, monkeys that experienced repeated housing changes early in development or were reared with peers rather than their mother showed a more rapid progression in indicators of virus activity. Furthermore, social reorganization stress at the time of viral inoculation was also associated with reduced survival. Moreover, recent studies from our group also observed that monkeys who progressed most rapidly following inoculation with SIV had reduced activity of cytotoxic T-cell (CD16+ lymphocytes) function before exposure to the virus and independent of early experiences. Although, many potential influences of early experiences on both immunoregulation and disease resistance in adult monkeys have been demonstrated, recent studies also indicate that the risk for rapid progression may rest in genetically determined aspects of immunoregulation as well.

Social Relationships Affect Immunoregulation

We have learned the most regarding the relation between behavior and immunity from observations of the negative consequences of stressors. Social psychology has shown for many years that the presence of social support in humans has positive consequences for health and well-being. There is growing evidence in nonhuman primates that a number of aspects of social relationships can also have positive consequences for markers of health and immunity. Most nonhuman primates are social creatures with their lives organized around social interactions with other members of their group. During times of social challenge, the presence of another monkey is important for reducing the negative endocrine and immune

consequences of the stressor. The negative consequences of maternal absence on behavior and physiology are significantly reduced in the presence of other monkeys who interact with the separated monkey. These social interactions may include affiliative behaviors, such as sitting nearby, play, and/or grooming. Being groomed by another monkey is quite effective in reducing heart rate to levels comparable to those noted during sleep. Grooming in monkeys strengthens social ties and facilitates reconciliation following fights between monkeys. Heart rate recovers rapidly following an aggressive interaction if one of the monkeys is groomed by the other involved in the fight. The immediate immune consequences of these behaviors have not been determined.

The relationship between dominance status in nonhuman primates and immune regulation is not as clear as we might expect. A high social rank does not necessarily imply enhanced immune capacity. Immune status assessed by *in vitro* methods in monkeys with high social status does not differ from low-ranking monkeys under stable social conditions nor do acute short-term stressors modulate these measures in a manner predicted by social status in the group. However, if the specific antibody responses to an antigen, an *in vivo* measure, are assessed under stable conditions, low-ranking monkeys mount a larger response than high-ranking monkeys. High rank may carry with it some subtle negative consequences. It appears that specific affiliative behaviors are most critical in modulating the impact of stressful social conditions in nonhuman primates, not unlike the role of social support for humans. Furthermore, the assumption that stress suppresses immune regulation is no longer entirely correct because acute stressors and challenges can enhance aspects of the inflammatory response.

Individual Differences in Immune Regulation

Monkeys differ along a number of behavioral and physiological dimensions. The concept of individual differences has received more attention from researchers in PNI. In the face of considerable variability in outcome following seemingly similar challenges, we may observe a wide range in the levels of immune measures even under resting conditions. Epidemiological research has tended to focus on demographic characteristics such as gender, age, socioeconomic status (SES), race, and so on in partitioning outcome variability. The long-term consequences of these factors have received considerable attention in a concept called allostatic load, the cumulative effects of long-term stressors that ultimately push the adaptational capacity of the stress response system beyond its capacity. However, there are individual factors that contribute to the extent to which these stressors affect immunoregulation.

Recent observations in monkeys have suggested that behavioral characteristics or traits unique to each monkey are influential in determining these outcomes. For example, sociability, or the tendency to engage in affiliative interactions with other monkeys, is associated with more rapid clearance of the virus in SIV-inoculated monkeys. These findings support the observations that social affiliation reduces the negative consequences of social challenge. Under free-ranging conditions, monkeys characterized as irritable (e.g., active, irritable, and aggressive) tended to have higher active infection with cytomegalovirus (CMV), a virus common in these animals. Thus, in addition to demographic characteristics, individual differences in behavioral traits also contribute to immunological variance. Other aspects of traitlike behaviors on nonhuman primates (e.g., personality) affect immune regulation to the extent that animals' relative sociality in social groups affect their specific antibody response following relocation. Monkeys rated as low in sociability (e.g., motivation to seek social interactions with others) had reduced specific antibody responses during this challenge.

Environmental Factors Influencing Immunoregulation

The complex nature of nonhuman primates presents the researcher with a number of challenges to provide for their psychological well-being. These manipulations are strictly monitored by federal (U.S. Department of Agriculture, USDA) and private (Association for Assessment of Laboratory Animal Care, AALAC) agencies. The simplest way of achieving this goal is through social housing and avoiding single-individual housing of nonhuman primates. Only a limited number of studies have actually investigated the impact of environmental enrichment on immune responses *per se*. A large number of studies, however, have investigated the impact of environmental enrichment on social behaviors of macaque monkeys. The provision of foraging enrichments (e.g., foraging boards and small preferred food treats mixed into the bedding) and manipulative devices such as simple toys significantly reduces the number of stereotypical behaviors and increase positive social interactions in nonhuman primates. As mentioned earlier, unpredictable housing experiences are associated with immune dysregulation as reflected in an attenuated specific antibody responses. Observations in SIV-infected macaques also indicate that the number of cage changes

occurring in advance of inoculation also impact the response to viral inoculation; that is, numerous cage moves, which represent a significant social stressor, are associated with more rapid progression and fatality to the virus. Features of housing for research animals that were neglected in the past have also been shown to impact experimental outcomes in significant ways.

Mechanisms of Brain–Immune Interactions

If behavior (i.e., the central nervous system) is likely to affect the regulation of the immune response, there must be potential pathways for these interactions. Indeed, there are two major means of brain and immune communication: direct neural innervation of immune organs by the autonomic nervous system and soluble blood-borne factors (i.e., cytokines and endocrine hormones). Nerve fibers of the autonomic nervous system come in close proximity to immune cells in lymphoid organs such as the spleen, thymus, and lymph nodes. Stress hormones such as cortisol are present in the blood during times of environmental and social challenge. Both of these routes, activated during stress, provide the means for neural–immune interactions. In addition, the cells of the immune system possess receptors for the neural and hormonal systems that permit the various neurotransmitters and hormones to affect immune regulation.

This is not a one-way street; the actions and products of the immune system itself also influence the central nervous system. A major component of illness is the associated changes that include sickness behaviors such as fatigue, sleep changes, changes in cognition, and loss of appetite. In many respects, these behaviors associated with cytokine release parallel major depressive disorders, and it has been argued that sickness behavior is one of many factors that underlie depression in clinical populations. Only a few studies have approached sickness behavior in primate models, but cytokines such as interleukin (IL)-1 and interferon-α have been observed to produce sleeplike behavior, advance rapid eye movement (REM) onset, and activate the HPA axis. These effects are consistent with observations in rodent models of sickness behavior and suggest that similar relationships to behavior exist in nonhuman primates as well.

Immune Senescence in Nonhuman Primates

It is generally assumed that aging is associated with immune downregulation. However, observations in successfully aging seniors have indicated that this may not universally be the case. As a group, older monkeys demonstrate reduced *in vitro* immune responses, such as lymphocyte proliferation and natural cytotoxicity. Functional measures indicate that, compared to a younger monkey, the older monkey's response to challenge with a trivalent flu vaccine is impaired. The extent to which cytotoxic responses are maintained with aging is an important predictor of long-term survival. It has been pointed out that the selection of old animals from a colony creates an inherent bias because they are really only members of a survivor cohort. In general, evidence supports the natural cytotoxic response as a potential proxy marker for successful aging and viral suppression in both old and younger animals.

Interestingly, feeding and exercise habits contribute significantly to the relationship between aging and immune senescence. For almost a century, it has been recognized that caloric restriction (not to be confused with malnutrition) is associated with a number of health benefits, including reduced morbidity and mortality in rodent models. A study that began in the late 1980s fed two groups of rhesus monkeys nutritionally fortified but either calorically restricted or *ad libitum* diets. As these monkeys now approach their senior years, a number of studies are underway to evaluate the impact of this restricted diet on brain structure, neuroendocrine regulation, and immune competence. Immune assessments have indicated attenuated age-related changes among the calorically restricted monkeys in many parameters, including *in vitro* cellular proliferation, cytokine levels, *in vitro* cytokine production, and cytokine gene expression, most notably in interferon-γ. These initial results suggest that some of the untoward aspects of aging are potentially attenuated in association with dietary restriction.

Summary

Nonhuman primates provide excellent models for research in PNI. The complexities of their behavioral interactions serve as an outstanding model for behavior and immune interactions relevant to humans. Early experiences affect immune regulation beyond the developmental period. Negative events experienced early in development contribute to increased risk for disease; however, social factors can reduce the impact of many of these negative experiences. Individual differences in behavioral characteristics also contribute to differences in immune regulation. There are abundant means whereby the immune system and the brain are able to interact. Understanding of these interactions poses a significant challenge for PNI in the twenty-first century.

Further Reading

Ader, R., Felten, D. L. and Cohen, N. (eds.) (2001). *Psychoneuroimmunology* (3rd edn.) San Diego, CA: Academic Press.

Capitanio, J. P. and Lerche, N. W. (1998). Social separation, housing relocation, and survival in simian AIDS: a retrospective analysis. *Psychosomatic Medicine* **60**, 235–244.

Coe, C. L. and Lubach, G. R. (2005). Prenatal origins in individual variation in behavior and immunity. *Neuroscience and Biobehavioral Reviews* **29**, 39–49.

Fleshner, M. and Laudenslager, M. L. (2004). Psychoneuroimmunology: then and now. *Behavioral and Cognitive Neuroscience Reviews* **3**, 114–130.

Glaser, R. and Kiecolt-Glaser, J. K. (2005). Stress-induced immune dysfunction: implications for health. *Nature Reviews: Immunology* **5**, 243–251.

Johnson, R. W. (2002). The concept of sickness behavior: a brief chronological account of four key discoveries. *Veterinary Immunology and Immunopathology* **8**, 443–450.

Kemeny, M. and Laudenslager, M. L. (eds.) (1999). *Special issue on individual differences. Brain, Behavior, and Immunity* **13**(2).

McEwen, B. S. (1998). Stress, adaptation, and disease: allostasis and allostatic load. *Annals of the New York Academy of Sciences* **840**, 33–44.

Roitt, I., Brostoff, J. and Male, D. (2001). *Immunology* (6th edn.). Philadelphia: Mosby.

Roth, G. S., Mattison, J. A., Ottinger, M. A., et al. (2004). Aging in rhesus monkeys: relevance to health intervention. *Science* **205**, 1423–1426.

Schapiro, S. J. (2002). Effects of social manipulations and environmental enrichment on behavior and cell-mediated immune responses in rhesus macaques. *Pharmacology, Biochemistry, and Behavior* **73**, 271–278.

Thierry, B., Singh, M. and Kaumanns, W. (eds.) (2004). *Macaque societies: a model for the study of social organization*. Cambridge, UK: Cambridge University Press.

Primate Models, Cardiovascular Disease

J R Kaplan
Wake Forest University School of Medicine, Winston Salem, NC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J R Kaplan, volume 3, pp 230–235, © 2000, Elsevier Inc.

Nonhuman Primates as Animal Models

Behavioral Influences on Risk Factors for Cardiovascular Disease

Behavioral Influences on Atherosclerosis

Behaviorally Induced Abnormalities in Vascular Reactivity

Summary

Glossary

Abnormal coronary vascular reactivity	The tendency for coronary arteries to lose their ability to dilate rather than constrict as demand for blood is increased.
Atherosclerosis	A gruel-like accumulation of fatty substance in the intima (inner layer) of the large muscular and elastic arteries of the body.
Beta-blocker	A class of drugs that inhibits the activation of the sympathetic nervous system (i.e., heart rate, blood pressure, and cardiac output) by blocking the response of β-adrenergic receptors in the arteries and heart.
Coronary heart disease (CHD)	A disease spectrum subsuming myocardial infarction (heart attack), angina pectoris (ischemic chest pain), and disruptions to the rhythm and electrical activity of the heart.
Dominant/subordinate social status	The position that an animal occupies in a pecking order. High-status animals (dominants) predictably defeat low-status animals (subordinates) in competitive encounters. Subordinate animals generally exhibit signs of stress (e.g., exaggerated release of cortisol from the adrenal cortex and behavioral withdrawal).
Endothelium	The inner wall of the artery; a normally functioning endothelium causes the artery to dilate and constrict appropriately in response to blood flow.
Heart rate reactivity	The degree to which an individual's heart rate increases in response to stressful situations. High heart rate reactivity is probably due to excessive activation of the sympathetic nervous system, and it is associated with the increased development of atherosclerosis in people and monkeys.
Hypertension	A sustained increase in blood pressure.

Ischemia	A reduction in blood flow to an organ or tissue, due to obstruction (e.g., a blood clot), arterial constriction, atherosclerosis, or a combination of factors.
Plasma lipids	Fats or fatlike substances (insoluble in water) that include total plasma cholesterol (TPC), high-density lipoprotein cholesterol (HDLC), and low-density lipoprotein cholesterol (LDLC). In general, TPC = HDLC + LDLC. Elevations in plasma levels of LDLC (and the resulting increases in TPC) are highly atherogenic. High plasma concentrations of HDLC (the good cholesterol) are protective.
Quantitative coronary angiography	An imaging procedure in which an opaque dye is infused into the coronary arteries, allowing them to be visualized radiographically and measured.
Sympathetic nervous system (SNS)	That part of the involuntary or self-regulating nervous system that controls responses (e.g., heart rate and blood pressure increases) to stressful situations.

Cardiovascular disease is manifested most prominently in human beings as coronary heart disease (CHD) and stroke. Atherosclerosis in the coronary arteries is a usual cause of CHD, whereas atherosclerosis in the cerebral arteries contributes to both transient ischemic attacks and ischemic stroke. Epidemiological and clinical studies have provided considerable evidence that behavioral factors influence the development of cardiovascular disease and the underlying atherosclerosis. However, logistical and ethical considerations prevent many questions regarding risk factors and mechanisms from being addressed directly in people. Animal models offer an alternative means of answering such questions.

Nonhuman Primates as Animal Models

The ideal animal for evaluating behavioral influences on cardiovascular disease would respond to a fat- and cholesterol-containing diet by exhibiting both an intimal accumulation of atherosclerosis in the large proximal coronary arteries and a susceptibility to infarction and ischemia. Furthermore, this animal would resemble human beings in cardiovascular and reproductive physiology and in neurobiology. Finally, the animal's behavioral responses and characteristics would vary along dimensions (e.g., aggressiveness and competitiveness, social support) analogous to those thought important in the epidemiology of atherosclerosis and CHD in human beings. Of all animal models, the nonhuman primates best approximate the foregoing characteristics.

Numerous investigators have used primates to explore one facet or another of the hypothesis that behaviorally evoked, excessive perturbations of the body's principal axes of neuroendocrine response (i.e., the adrenal cortex or the SNS) may produce pathophysiological consequences. Some of these studies provide data regarding the influence of stress on the risk factors for cardiovascular disease (e.g., elevations in plasma lipids and blood pressure) or on mediating factors (circulating concentrations of immune or inflammatory markers), whereas others have focused primarily on the characteristics of atherosclerotic lesions or abnormalities in vascular reactivity to changes in blood flow. Taken together, existing studies present a clear and convincing case that behavioral factors (either stressful environments or particular characteristics of temperament) influence both atherosclerosis and clinical events in ways that have relevance for human beings.

Behavioral Influences on Risk Factors for Cardiovascular Disease

Plasma Lipids

Elevations in plasma lipids (e.g., total plasma cholesterol, TPC; low-density lipoprotein cholesterol, LDLC) are associated with the exacerbation of atherosclerosis and increased risk of CHD and ischemic stroke. Thus, behavioral factors may contribute to cardiovascular disease through a mechanism involving adverse changes in plasma lipids. Among primates, the lipids of squirrel monkeys (*Saimiri sciurius,* a New World primate) seem particularly susceptible to behavioral influences. In one study, for example, there was an increase in plasma lipids as animals were transported from South America to a laboratory in the United States. Interestingly, both TPC and LDLC declined in these animals after they were put in a secluded environment, but increased again following exposure to a busy laboratory. All of these changes occurred in the absence of any cholesterol in the diet and are attributed most reasonably to the influence of psychic stress. Cholesterol-fed squirrel monkeys are similarly responsive to experimental manipulations; exposure to either a shock-avoidance paradigm (in which the animal presses a lever to avoid a shock) or a restraint box leads to a significant and sustained increase in TPC and LDLC.

Among the Old World monkeys, adverse serum lipid profiles have been observed in relation to individual differences in social status and individual differences in heart rate reactivity to stress. In both baboons (*Papio hamadrayas*) and cynomolgus macaques

(*Macaca fascicularis*), for example, subordinate social status is associated with suppressed plasma concentrations of high-density lipoprotein cholesterol (HDLC), which is a risk factor for atherosclerosis. Notably, cynomolgus macaques that reliably exhibit exaggerated heart rate responses to psychological stress (e.g., exposure to a threatening object) also tend to have suppressed plasma concentrations of HDLC.

Data from Old and New World monkeys suggest that stressful environments, low social status, and exaggerated SNS responsivity to challenge are all associated with potentially harmful plasma lipid profiles. Among human beings, there are also numerous observations relating either elevations in LDLC or reductions in HDLC to stressful circumstances or particular personality characteristics.

Blood Pressure

Elevated blood pressure (hypertension) is associated with increased risk of CHD and stroke. High blood pressure also exacerbates atherosclerosis in the coronary and cerebral arteries. Numerous studies indicate that behavioral stress can lead to sustained, potentially pathogenic increases in both systolic and diastolic blood pressure. For example, squirrel monkeys trained to avoid a shock by pressing a lever ultimately exhibit increases in blood pressure that persist across experimental sessions. Rhesus monkeys (*M. mulatta*) exposed to a similar manipulation also develop a steady elevation in blood pressure. Further, rhesus monkeys as well as baboons experience sustained increases in blood pressure when light signals and shock are paired and combined with a bar press for the delivery of food (called conflict manipulation). Chronic exposure to loud noise also produces elevated blood pressure in cynomolgus and rhesus macaques. Notably, adult male baboons that typically live together with a number of females in a harem develop hypertension when they are separated from their harems and forced to watch the females feed first. Males also exhibit a hypertensive response when they are separated from their harems and are forced to observe a new male being introduced to these females. These latter manipulations are purely social and more closely model the kind of stress most likely to affect human beings.

Behavioral Influences on Atherosclerosis

Atherosclerosis in Males

Primates typically do not develop atherosclerosis to any significant degree unless they are fed a diet relatively high in saturated fat and cholesterol, similar to the diet often consumed by people living in industrialized nations. Early investigations used physical restraint or exposure to a shock-avoidance manipulation to provide evidence that psychic stress, in combination with an atherogenic diet, could potentiate atherosclerosis in monkeys. However, social manipulations also influence the development of atherosclerosis in monkeys. For example, when cynomolgus males are housed in groups that either remain stable in composition or that are perturbed (by distributing members across groups), dominant and subordinate animals tend to retain their relative rankings irrespective of social condition (i.e., stable or unstable groups). A quantitative evaluation of the coronary arteries of such animals revealed that monkeys housed in an unstable condition had significantly more coronary artery atherosclerosis than animals in a stable environment but only if they were dominant in their groups. This finding suggests that the behavioral demands of retaining dominant social status in an unstable social setting potentiate atherosclerosis, at least in males of this species consuming an atherogenic diet. A similar pattern was obtained (although with markedly smaller lesions), among animals that consumed a diet modeled on the recommendations of the American Heart Association. Thus, behavioral factors might, even in the absence of dietary provocation, cause damage that could initiate the earliest lesions of atherosclerosis. Subsequent investigations have shown that exposure to the stress of social perturbation for as little as 72 h can cause the degeneration and death of endothelial cells, thus marking the first stage of atherosclerosis progression.

Mediating Role of Sympathetic Nervous System Activation in the Atherosclerosis of Males

Significantly, behavioral effects observed in the social manipulation experiments were independent of variation in serum lipid concentrations and resting blood pressure. Numerous investigators have proposed that recurrent SNS activation in response to behavioral stimulation could promote the development of atherosclerosis and its clinical sequelae. Testing the hypothesis that excessive SNS activation might have potentiated atherogenesis among dominant animals in the foregoing studies, male monkeys in a subsequent investigation were all housed in unstable social groups and fed an atherogenic diet; here, half of the monkeys also received a beta-blocker, propranolol, throughout the 2-year study. Untreated dominant monkeys again had twice the amount of atherosclerosis of their subordinate counterparts,

replicating the results from the unstable treatment condition of earlier experiments. Dominant monkeys treated with propranolol, however, developed less than half the coronary atherosclerosis of untreated dominants and had roughly the same extent of lesion as both treated and untreated subordinate animals. Hence, the more pronounced atherosclerosis characteristic of untreated (i.e., autonomically intact) dominant monkeys, when housed in groups of rotating membership and fed a cholesterol containing diet, can be prevented by the long-term administration of a beta-blocker and is therefore dependent on concomitant SNS activation. Furthermore, the monkey studies also provide evidence that behaviorally induced SNS activation initiates atherogenesis independently of the presence of a high-fat diet. Hence, the endothelial cell death associated with the short-term (72-h) perturbation previously described occurs in the absence of elevated plasma lipids and can be inhibited by pretreatment with a beta-blocker.

Other data from this species further support the speculation that a heightened SNS responsivity to behavioral stimuli increases the risk for atherosclerosis. Specifically, cynomolgus macaques display reliable and significant individual differences in heart rate response (assessed via radiotelemetry) to a standardized social stressor (exposure to an unknown human). Animals can thus be characterized as either high or low in heart rate reactivity to stress. Notably, the atherosclerotic lesions of high heart rate reactors are twice as large as those of their low-reactive counterparts. Taken together, the results of the macaque experiments indicate that some individuals (e.g., high heart rate reactors) may be susceptible to the development of atherosclerosis because of an intrinsic SNS hyperresponsivity to behavioral stimuli, whereas others (dominant individuals in unstable social environments) may be at increased risk because of frequent elicitation of such responses in stressful environments. SNS activation, provoked in either of two ways, may thus provide the common pathway by which pathogenesis is mediated in male monkeys.

Summary for Male Monkeys

The studies of male monkeys relate a behavioral predisposition (social dominance, i.e., successful aggression) and excessive SNS stimulation to the exacerbation of atherosclerosis. Both domains appear to have analogs among human beings. There are other behavioral risk factors, however, about which the studies in male monkeys provide little data or insight. Such factors include depression, sleep disturbances, and a propensity to anxiety or panic, all of which have been associated reliably with elevated risk of CHD in human beings. Moreover, social support is often believed to be cardioprotective. Although several measures of social support have been evaluated in the foregoing studies of male monkeys, there is no evidence that animals differing by these measures vary in their extent of atherosclerosis.

Sex Differences in Atherosclerosis

The relative sparing of premenopausal women compared to men of similar ages is a prominent feature of CHD, ischemic stroke, and atherosclerosis. Although this phenomenon is sometimes referred to as female protection, it is characterized more accurately as a delay in disease onset, with the incidence curve for women lagging behind that of men by approximately 10 years. The various effects of estrogen are believed to account for most of this sex difference, at least with respect to CHD incidence and underlying atherosclerosis. Not only are premenopausal women protected from CHD, but the provision of estrogen replacement to women immediately prior to or immediately following menopause is associated with a significant reduction in CHD risk (it must be noted that such protection is not apparent when hormone therapy is initiated in women with preexisting disease or who are 10 or more years postmenopausal).

Because atherosclerosis progresses over decades, it is likely that the clinical events occurring in postmenopausal women have their beginnings in the premenopausal years. This conclusion is supported by autopsy assessments showing that by 34 years of age, at least one-third of all women (regardless of race) have raised lesions in their coronary arteries. *In vivo* arterial imaging also demonstrates the presence of relatively extensive focal atherosclerosis in premenopausal women. The observations that atherosclerosis begins premenopausally raises the question of whether there are particular risk factors that place some women on a higher-risk trajectory than others. Considerable effort has been focused on this question in studies employing premenopausal monkeys as a model of reproductive-age women. The major results of these studies are described next.

Social Stress and Female Protection from Atherosclerosis in Monkeys

Investigations using socially housed female cynomolgus monkeys provide evidence that stress-induced ovarian dysfunction eliminates premenopausal protection from atherosclerosis. For example, in one 2-year experiment contrasting the atherosclerosis of socially housed males and females, the females were divided

into groups of five. Each of these groups contained a vasectomized male; additional males were placed together in groups of five. All animals consumed an atherogenic diet. Quantitative evaluation of the coronary arteries at the end of this experiment demonstrated that female protection with regard to atherosclerosis extended only to the dominant animals (i.e., those ranking first or second in their social groups); subordinate females (ranking third or below) were indistinguishable from males. Subordinate females also had five times as many anovulatory cycles and three times as many cycles characterized by luteal-phase progesterone deficiencies as did their dominant counterparts. Notably, females with the most extensive atherosclerosis were all subordinate, and all had marked ovarian endocrine dysfunction. These data suggest that the stress of social subordination may increase the risk of atherosclerosis in females by inducing a relative ovarian endocrine deficiency state similar to that observed in postmenopausal women. Interestingly, investigations in premenopausal women also have also linked stress-induced estrogen deficiency to severity of angiographically documented CHD.

In a further study, ovariectomized females were placed in social groups for the same amount of time and fed the same diet as the reproductively intact females just described. The ovariectomized females as a group had significantly more extensive coronary artery atherosclerosis than the intact females. This effect was due entirely to the loss of protection experienced by the ovariectomized dominant animals compared with their intact counterparts. The subordinates, intact and ovariectomized, were affected equally with atherosclerosis. Additional studies have extended these findings by showing that the treatment of premenopausal females with either an oral contraceptive or unopposed estrogen treatment eliminates the difference between dominant and subordinate individuals. Parenthetically, it should be noted that subordinate females tended to be hypercortisolemic compared to dominants, leading some scientists to suggest that the atherosclerosis distinctions between such animals may be mediated in part by the excessive activation of the hypothalamic-pituitary-adrenocortical axis. The fact that dominant and subordinate animals continue to differ in adrenocortical activation – but not atherosclerosis extent – following ovariectomy argues against this hypothesis.

Finally, a recent life-course study demonstrated that premenopausal behavioral and hormonal status predict postmenopausal atherosclerosis outcomes independently of any postmenopausal hormone exposure. Specifically, females that were subordinate in their social groups – and thus were ovarian-suppressed and estrogen-deficient – developed more atherosclerosis premenopausally and postmenopausally than did their dominant, ovarian-normal counterparts. Significantly, the worsened atherosclerosis of the subordinate, socially stressed individuals could be mitigated by premenopausal treatment with exogenous estrogen in the form of an oral contraceptive. This important set of observations implies for women that premenopausal behavioral and associated hormonal (estrogenic) characteristics establish the trajectory of atherosclerosis progression that ultimately culminates in postmenopausal CHD events. The monkey research thus emphasizes the importance of early pathogenesis and early intervention, a perspective supported by recent reevaluations of clinical trial data.

The different patterns in which social status influences atherosclerosis in male and female monkeys deserves comment. Subordinate female primates, like subordinate and stressed females of many other mammalian species, tend to experience impaired ovarian function. This impairment eliminates female protection from atherosclerosis in monkeys and, by implication, women. Subordinate male monkeys also experience stress, possibly including impaired reproductive activity; however, the physiological sequelae of social subordination among males do not seem to affect the coronary arteries. Instead, dominant males in unstable settings are predisposed to atherosclerosis because they experience marked elevations in artery-damaging SNS activity as they attempt to maintain behavioral preeminence in the face of challenge.

Social Isolation, Depression, and Atherosclerosis in Females

The exacerbation of atherosclerosis in female monkeys is also observed in conjunction with single-cage housing (i.e., social separation). Specifically, females housed in single cages have significantly more coronary artery atherosclerosis than comparably treated animals housed in social groups. Single-caged animals are most similar in atherosclerosis extent to socially housed subordinates. Interestingly, this worsening of atherosclerosis among single-caged females is not related to ovarian impairment; rather, such animals exhibit elevations in heart rate. This latter observation suggests that, at least in female monkeys, social separation exacerbates coronary artery atherosclerosis via exaggerated SNS activity. No similar studies have been done in male monkeys. This outcome in females is, however, consistent with extensive literature documenting the cardioprotective

effects of social support and the detrimental impact on CHD of social isolation in human beings.

Epidemiological studies have linked depression to CHD, and it is well known that women suffer from a higher incidence of depression than men. The role of depression in the coronary artery atherosclerosis of female monkeys has not been systematically evaluated, in part because primate depression usually has been studied with respect to developmental abnormalities caused by early separation of infants from their mothers. Recently, however, it has been shown that a proportion of socially housed, premenopausal cynomolgus monkeys exhibited behavioral and physiological characteristics that model adult human depression. This model offers for the first time the possibility of determining experimentally whether depression is causally associated with the development of atherosclerosis, the primary underlying cause of CHD.

Behaviorally Induced Abnormalities in Vascular Reactivity

Evaluation of Abnormalities in Vascular Reactivity

Numerous investigators have proposed that behavioral stress contributes to the occurrence of myocardial ischemia by causing the arteries to constrict inappropriately, thus cutting off the flow of blood to the heart. The potential for ischemia is evaluated by assessing the responses of coronary arteries to the infusion of vasoactive substances in the context of behavioral investigations. The normal endothelium, for example, dilates in response to acetylcholine. In contrast, endothelial damage causes a paradoxical constriction in response to acetylcholine. Nitroglycerine, an endothelium-independent dilator, reliably causes arterial dilation irrespective of the condition of the endothelium. Using such vasoactive substances and appropriate imaging techniques, it is now possible to visualize and quantitate the responsivity of the coronary arteries of monkeys exposed to various kinds of stress. Such functional assessments have markedly extended the usefulness of monkeys in cardiovascular research.

Studies in Male Monkeys

In one such study, quantitative coronary angiography was used to evaluate arterial responses to acetylcholine and nitroglycerine in male cynomolgus monkeys that either were or were not exposed to the stress of repeated reorganization of social-group membership (i.e., social instability). The coronary arteries of animals living in unstable groups constricted in response to acetylcholine, whereas those of monkeys from stable groups dilated. Significantly, the degree of constriction observed in this experiment was independent of the extent of the underlying atherosclerosis; animals with little atherosclerosis but living in unstable groups had a degree of constriction equivalent to unstable monkeys that had extensive atherosclerosis. In contrast, monkeys living in stable groups dilated, irrespective of the amount of atherosclerosis.

The previous study, and others like it, demonstrate that chronic social stress impairs endothelium-dependent vascular responses, an effect that is not necessarily dependent on the extent of atherosclerosis. It should be noted that none of these investigations revealed an influence of dominance status on vascular reactivity. Hence, the pattern of behavioral influences on blood flow (and thus on ischemia) differs somewhat from that observed with respect to atherosclerosis, at least among socially housed cynomolgus males.

Studies in Female Monkeys

As with atherosclerosis, behavioral influences on vascular reactivity among females are apparently mediated by social subordination and ovarian dysfunction. Hence, the coronary arteries of socially subordinate female cynomolgus monkeys, compared to dominants, constrict in response to acetylcholine. In contrast, nitroglycerine dilates the arteries of all monkeys, irrespective of status. As in the other studies, subordinates here were characterized by ovarian dysfunction (low plasma estradiol and luteal-phase abnormalities in progesterone). Notably, the extent of vascular constriction in response to acetylcholine was directly related to the degree of ovarian impairment as indicated by plasma estradiol concentrations on the day of testing. These observations not only demonstrate the role of behavioral factors on coronary vascular reactivity among females, they emphasize again the cardioprotective significance of normal ovarian function.

Summary

Systematic studies in nonhuman primates have shown that stress, both social and nonsocial, consistently affects the risk factors for CHD and the underlying atherosclerosis. Studies involving socially housed individuals are probably of most relevance to human beings; such investigations have focused primarily on the development of coronary artery atherosclerosis in macaque monkeys. Stress clearly exacerbates atherosclerosis in these animals, albeit differently in males and females.

Among males, dominant individuals in unstable settings and those who exhibit exaggerated heart rate reactivity to challenge are at greatest risk. These outcomes are most likely mediated by excessive SNS activation. These findings are reminiscent of observations in men linking excessive hostility and high heart rate reactivity to increased CHD risk. The monkey results imply that the pathogenicity of hostility (approximated in the animals by social dominance) is most prominent under stressful disrupted social conditions. Notably, indices of social support have not been found to predict patterns of atherosclerosis among male monkeys.

The situation among females is different. Normal ovarian function seems critical for the premenopausal inhibition of atherosclerosis among animals consuming a diet relatively high in fat and cholesterol. Notably, among groups of females, animals occupying the lower status positions are stressed and experience ovarian impairment and relative estrogen deficiency. Such individuals are characterized by a rapid acceleration of premenopausal atherosclerosis, which can be inhibited by exogenous estrogen. A novel observation is that the atherosclerosis trajectory established by premenopausal behavioral and hormonal conditions predicts the postmenopausal atherosclerosis extent. These findings suggest that primary prevention of CHD in women should begin prior to or immediately at menopause. The studies in female monkeys also identify social isolation and social support as potential factors affecting disease risk. Systematic studies of the role played by depression in the development of atherosclerosis have not yet been conducted, but they hold promise for the future.

Further Reading

Bairey Merz, C. N., Johnson, B. D., Sharaf, B. L., et al. (2003). Hypoestrogenemia of hypothalamic origin and coronary artery disease in premenopausal women: a report from the NHLBI-sponsored WISE study. *Journal of the American College of Cardiology* **41**, 413–419.

Cohen, S., Kaplan, J. R. and Manuck, S. B. (1994). Social support and coronary heart disease: underlying psychologic and biologic mechanisms. In: Schumaker, S. & Czajkowski, S. M. (eds.) *Social support and cardiovascular disease*, pp. 195–221. New York: Plenum Press.

Henry, J. P. and Stephens, P. M. (1977). *Stress, health and the social environment: a sociobiologic approach to medicine*. New York: Springer-Verlag.

Kaplan, J. R. and Manuck, S. B. (1999). Status, stress, and atherosclerosis: the role of environment and individual behavior. *Annals of the New York Academy of Sciences* **896**, 145–161.

Kaplan, J. R. and Manuck, S. B. (2004). Ovarian dysfunction, stress, and disease: a primate continuum. *ILAR Journal* **45**, 89–115.

Kaplan, J. R., Manuck, S. B., Anthony, M. S., et al. (2002). Premenopausal social status and hormone exposure predict postmenopausal atherosclerosis in female monkeys. *Obstetrics and Gynecology* **99**(3), 381–388.

Kaplan, J. R., Manuck, S. B., Clarkson, T. B., et al. (1985). Animal models of behavioral influences on atherogenesis. *Advances in Behavioral Medicine* **1**, 115–163.

Klaiber, E. L., Vogel, W. and Rako, S. (2005). A critique of the Women's Health Initiative hormone therapy study. *Fertility and Sterility* **84**, 1589–1601.

Mikkola, T. S. and Clarkson, T. B. (2006). Coronary heart disease and postmenopausal therapy: conundrum explained by timing? *Journal of Women's Health* **15**, 51–53.

Rozanski, A., Blumenthal, J. A. and Kaplan, J. R. (1999). Impact of psychological factors on the pathogenesis of cardiovascular disease and implications for therapy. *Circulation* **99**, 2192–2217.

Schneiderman, N. (1987). Psychophysiologic factors in atherogenesis and coronary artery disease. *Circulation* **76**(supplement I), I41–I47.

Shively, C. S., Register, T. C., Friedman, D. P., et al. (2005). Social stress-associated depression in adult female cynomolgus monkeys (*Macaca fascicularis*). *Biological Psychology* **69**, 67–84.

Skantze, H. B., Kaplan, J., Pettersson, K., et al. (1998). Psychosocial stress causes endothelial injury in cynomolgus monkeys via α_1-adrenoceptor activation. *Atherosclerosis* **136**, 153–161.

Primate Models, Overview

D M Lyons
Stanford University Medical School, Stanford, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D M Lyons, volume 3, pp 236–240, © 2000, Elsevier Inc.

Social Psychological Sources of Stress
Psychobiological Consequences of Stress
Environmental and Genetic Risk Factors

Glossary

Model	Simplified experimental system used as a testing device.
Primates	Mammalian vertebrates with prehensile hands.
Social psychobiology	The study of relationships between social, psychological, and biological processes.

Practical limitations and ethical concerns restrict opportunities for prospective controlled studies of stress in healthy humans and patient populations. Animal models that use rodents are limited because corticolimbic brain substrates of social, emotional, and cognitive control of behavior differ significantly in rats and mice compared to humans and nonhuman primates. Studies of monkeys, baboons, and apes have clearly advanced our understanding of the social psychological sources of stress, psychobiological consequences of stress, and individual differences in vulnerability and resilience related to environmental and genetic risk factors.

Social Psychological Sources of Stress

Models of stress in rats and mice often rely on exposure to cold, electric shock, or physical restraint, whereas primate models highlight the significance of seemingly subtle sources of social psychological stress.

Maternal Separation and Loss

Primate infants generally respond to maternal separation with increased secretion of hypothalamic-pituitary-adrenal axis stress hormones. The hypothalamic release of corticotropin releasing hormone (CRH) stimulates pituitary secretion of adrenocorticotropic hormone (ACTH), which triggers the secretion of glucocorticoids (i.e., cortisol) from the adrenal cortex. An informative exception to the rule that maternal separation elicits a robust stress response is found in one of the few primate species in which fathers provide parental care. In titi monkeys, maternal separation does not elicit an adrenocortical stress response unless the father is also removed from the developing infant. Separation from the father invariably elicits a robust adrenocortical stress response even when infant titi monkeys are left in the company of their mother. These findings concur with behavioral evidence that fathers are the primary attachment figure for titi monkey infants.

Relationship Formation and Social Reunions

Among adult primates, the formation of new relationships is often stressful, as demonstrated by measures of social behavior, elevated plasma levels of cortisol, and evidence of impaired immunity. Even reunions with well-known relatives may elicit surprising effects. When previously separated adolescent mangabey monkeys are returned to their natal social group, these monkeys and their mothers engage in frequent affiliative behavior. Nevertheless, these reunions evoke in adolescent offspring increased plasma levels of cortisol and diminished cell-mediated immunity.

Crowding

High-density living conditions are especially stressful for certain individuals, as illustrated in rhesus macaques. A statistically significant interaction between crowding and behavioral inhibition (i.e., shyness) was found to predict individual differences in the incidence of physical injury. Compared to uninhibited individuals, highly inhibited cage-mates experienced higher rates of injury during (but not before or after) crowding in the presence of ample food. Evidently, inhibited rhesus macaques are targeted for violence in high-density conditions by uninhibited peers.

Competitive Conflicts and Foraging Demands

Stressful competitive conflicts occur during foraging and feeding in primates. Even in species in which overt aggression is uncommon, there is a general tendency toward mutual avoidance during foraging and feeding. Mutual avoidance minimizes fighting when food sources are dispersed, but when food is confined to limited locations stressful conflicts often emerge. Foraging demands likewise affect maternal behavior in vervet monkeys, squirrel monkeys, baboons, and macaques. Mother-reared bonnet macaque monkeys exposed as infants to stressful foraging demands

show increased emotional reactivity when challenged later in life by maternal separation or acute exposure to novelty. As adults, these bonnet macaque monkeys are also less social, more subordinate, and present with increased cerebrospinal fluid CRH levels.

Dominance Hierarchies and Social Conflicts

Primates that live in large groups can generally be ranked within a pecking order or social dominance hierarchy. The stressful aspects of hierarchies in primate societies are not necessarily related to a particular rank *per se*, but may reflect the style by which rank is expressed and the overall context in which social hierarchies emerge. In stable hierarchies, high-ranking primates tend to exhibit metabolic, immune, and cardiovascular profiles that appear to be less pathogenic than those that are seen in low-ranking subordinates. These rank-related differences are diminished during periods of recurring conflict and social instability. High-ranking primates in stable hierarchies also tend to have lower plasma levels of cortisol, but the levels maintained in dominant baboons that lack rudimentary social skills are as high as those in subordinates.

Psychobiological Consequences of Stress

Various psychobiological consequences of stress have been characterized in primate models at multiple levels of analysis.

Hippocampal Neuropathology

Cellular and molecular markers of hippocampal neuropathology have been identified in adult monkeys following exposure to chronic stress. Similar neuropathological damage is induced when sustained-release cortisol pellets are implanted in the hippocampus. Stress-induced reductions in glucocorticoid receptor binding sites within the hippocampus impair negative feedback sensitivity on dexamethasone suppression tests in cynomolgus macaques. Recent findings in New and Old World primates indicate that the hippocampal dentate gyrus continues to generate new neurons throughout adulthood and that neurogenesis is suppressed by exposure to psychosocial stress.

Neurochemistry and Cognitive Control

Squirrel monkeys that respond to social separation with greater increases in cortisol exhibit greater increases in cerebrospinal fluid levels of the dopamine metabolite homovanillic acid. Stress increases prefrontal dopamine release, and drugs that block dopamine receptors ameliorate stress-induced prefrontal dependent cognitive deficits in rhesus macaques. Stress also decreases serotonin in prefrontal brain regions and reduces cerebrospinal fluid levels of the serotonin metabolite 5-hydroxyindoleacetic acid (5-HIAA). Low levels of cerebrospinal fluid 5-HIAA are associated with a higher incidence of impulsive risk-taking behavior and excessive consumption of alcohol in adolescent macaques.

Reproductive Biology

Stress inhibits the release of hypothalamic-pituitary-gonadal axis hormones, alters the onset of puberty, impairs reproductive behavior, suppresses lactation, and adversely affects diverse aspects of parental care. The specific mechanisms that produce these outcomes are not well understood. A role for endogenous opioids is indicated by the observation that the stress-induced suppression of hypothalamic-pituitary-gonadal axis hormones is blocked by the administration of the opiate receptor antagonist naloxone. Other studies suggest that the stress-induced inhibition of reproduction occurs through nonopioidergic mechanisms and that hypothalamic-pituitary-adrenal axis hormones mediate these antireproductive effects.

Immunology

Stressful conflicts, the disruption of relationships, and the formation of new social groups impair immune responses in primates. Glucocorticoid levels within the range normally achieved in response to stress suppress numerous aspects of immunity, including lymphocyte numbers, lymphocyte function, and *in vivo* antibody responses to antigenic challenges. An alternative pathway for these immunological effects is through the catecholaminergic innervation of the lymph nodes, thymus, and spleen.

Health Disorders and Disease

Primate models provide evidence for adverse stress effects in coronary heart disease, susceptibility to upper respiratory infections, and vulnerability to simian acquired immune deficiency syndrome (the primate homolog of AIDS). Sustained exposure to stress induces metabolic abnormalities that parallel those found in humans who develop abdominal obesity and early-onset diabetes. Stress exacerbates the effects of high dietary salt intake on hypertension in baboons. Sustained exposure to stress elicits changes in primate behavior and neurobiology that resemble features of human anxiety disorders, alcoholism, and major depression.

Environmental and Genetic Risk Factors

Stress is an inescapable aspect of life that cannot be eliminated nor avoided for long. Equally obvious is that certain individuals are particularly vulnerable to

stress, whereas others seem unscathed. Experimental studies of primates indicate that environmental and genetic risk factors contribute to individual differences in vulnerability and resilience.

Prenatal Development

Pregnant squirrel monkeys and rhesus macaques exposed to chronic sources of stress produce infants with deficits in muscle tone, impaired neuromotor coordination, and shortened attention spans. Increased secretion of cortisol induced in mothers by the administration of exogenous ACTH during midgestation results in neonatal outcomes similar to those produced by prenatal stress. Later in life, prenatally stressed monkeys, and monkeys derived from ACTH-treated mothers, exhibit exaggerated neuroendocrine responses to stress and deficits in immune physiology.

Postnatal Development

Rhesus macaque infants raised without mothers show prolonged and inappropriate responses to stress, self-directed emotional behavior, diminished norepinephrine neurotransmission, altered autonomic reactivity, and impaired immune responses. Neuronal changes in dendritic branching have been characterized in motherless monkeys, and selective alterations in tyrosine hydroxylase activity are evident in the striatum, substantia nigra, and ventral tegmental area. Attempts to rehabilitate motherless monkeys may partially ameliorate anomalies in behavior but allegedly fail to normalize neurobiological effects.

Early-Life Stress Inoculation

Severely stressful early experiences increase the risk for the development of adult psychiatric conditions, including depression and anxiety disorders. Far less researched, but of equal importance, is the suggestion that mild forms of stress, instead of increasing vulnerability, result in subsequent stress resistance. Variously described as inoculating, immunizing, steeling, toughening, or thriving, the notion that mild early-life stress reduces vulnerability to subsequent stressors has important implications for understanding the neurobiology and prevention of mental health disorders. Recent studies of monkeys suggest that early exposure to mild stress leads subsequently to diminished measures of anxiety, pro-social tendencies, enhanced cognitive control, larger prefrontal brain volumes, and attenuated activation of the hypothalamic-pituitary-adrenal axis.

Social Buffering

The deleterious effects of stress on immune physiology are often buffered or reduced by the presence of familiar companions in primates with social proclivities. Under certain conditions, familiar companions likewise buffer or reduce the stress-induced secretion of glucocorticoids. Although glucocorticoids are known to regulate immune processes involved in susceptibility to disease, there is limited evidence in primates to implicate glucocorticoids as a direct mediator of social stress-buffering effects.

Phylogenetic Constraints

Species differences in neuroendocrine, autonomic, and behavioral responses to stress are important concomitants of species-typical modes of interacting with social and nonsocial environments. Interspecies differences in stress physiology and emotional vulnerability are well documented for two New World primates (squirrel monkeys and titi monkeys) and for at least five different species of macaques.

Inherited Variation

Coping with stress is increasingly recognized as key feature of health. Stimulated by the concept of primary prevention, health professionals are developing programs aimed at teaching people how to cope effectively with stress. That modes of coping are learned and not inherited from others is implicit in this approach. Yet human twin studies clearly suggest robust hereditary influences on aspects of coping with stress. Individual differences in autonomic reactivity to stress and cerebrospinal fluid levels of neurotransmitter metabolites are likewise heritable in rhesus macaques. High heritabilities for total brain size have been reported in humans and monkeys, but little is know about the genetics of inherited variation in stress-related corticolimbic brain circuits.

See Also the Following Articles

Animal Models (Nonprimate) for Human Stress; Crowding Stress; Hippocampus, Overview; Hypothalamic-Pituitary-Adrenal; Primate Hierarchies and Personality; Primate Models, Behavioral-Immunological Interactions; Primate Models, Cardiovascular Disease; Psychosocial Factors and Stress; Reproductive Dysfunction in Primates, Behaviorally Induced; Social Stress, Animal Models of.

Further Reading

Abbott, D. H., Keverne, E. B., Bercovitch, F. B., et al. (2003). Are subordinates always stressed?: a comparative analysis of rank differences in cortisol levels among primates. *Hormones and Behavior* **43**, 67–82.

Arnsten, A. F. (2000). Stress impairs prefrontal cortical function in rats and monkeys: role of dopamine D1 and norepinephrine alpha-1 receptor mechanisms. *Progress in Brain Research* **126**, 183–192.

Bailey, M. T., Lubach, G. R. and Coe, C. L. (2004). Prenatal stress alters bacterial colonization of the gut in infant monkeys. *Journal of Pediatric Gastroenterology and Nutrition* **38**, 414–421.

Barr, C. S., Schwandt, M. L., Newman, T. K., et al. (2004). The use of adolescent nonhuman primates to model human alcohol intake: neurobiological, genetic, and psychological variables. *Annals of the New York Academy of Sciences* **1021**, 221–233.

Boyce, W. T., O'Neill-Wagner, P., Price, C. S., et al. (1998). Crowding stress and violent injuries among behaviorally inhibited rhesus macaques. *Health Psychology* **17**, 285–289.

Gorman, J. M., Mathew, S. and Coplan, J. (2002). Neurobiology of early life stress: nonhuman primate models. *Seminars in Clinical Neuropsychiatry* **7**, 96–103.

Gould, E. and Gross, C. G. (2002). Neurogenesis in adult mammals: some progress and problems. *Journal of Neuroscience* **22**, 619–623.

Hoffman, K. A., Mendoza, S. P., Hennessy, M. B., et al. (1995). Responses of infant titi monkeys. *Callicebus moloch,* to removal of one or both parents: evidence for paternal attachment *Developmental Psychobiology* **28**, 399–407.

Kalin, N. H. (2004). Studying non-human primates: a gateway to understanding anxiety disorders. *Psychopharmacology Bulletin* **38**, 8–13.

Kaplan, J. R. and Manuck, S. B. (2004). Ovarian dysfunction, stress, and disease: a primate continuum. *ILAR Journal* **45**, 89–115.

Lyons, D. M. (2002). Stress, depression, and inherited variation in primate hippocampal and prefrontal brain development. *Psychopharmacology Bulletin* **36**, 27–43.

Parker, K. J., Buckmaster, C. L., Schatzberg, A. F., et al. (2004). Prospective investigation of stress inoculation in young monkeys. *Archives of General Psychiatry* **61**, 933–941.

Schneider, M. L., Moore, C. F., Kraemer, G. W., et al. (2002). The impact of prenatal stress, fetal alcohol exposure, or both on development: perspectives from a primate model. *Psychoneuroendocrinology* **27**, 285–298.

Shively, C. A., Register, T. C., Friedman, D. P., et al. (2005). Social stress-associated depression in adult female cynomolgus monkeys (*Macaca fascicularis*). *Biological Psychology* **69**, 67–84.

Tilbrook, A. J., Turner, A. I. and Clarke, I. J. (2002). Stress and reproduction: central mechanisms and sex differences in non-rodent species. *Stress* **5**, 83–100.

Primates: Rearing and Effects of Stress on Primate CNS Function

T K Newman
University of Cape Town, Cape Town, South Africa
C S Barr
National Institutes of Health, Bethesda, Maryland, USA

© 2007 Elsevier Inc. All rights reserved.

Nonhuman Primate Models of Early Stress
Gene by Environment Interaction

Glossary

Gene by environment interaction — Describes how a phenotype (e.g., impulsive behavior) is the result of both genetic and non-genetic (environmental) factors such that the influence of an environmental effect is conditioned by the individual's genotype.

Hypothalamic-pituitary-adrenal (HPA) axis — The hypothalamus, pituitary gland and adrenal cortices form a linked neuroendocrine system (axis) that controls an organism's reaction to stress through the secretion and regulation of a hormonal cascade. Activation of the HPA axis induces the release of CRH (corticotropin-releasing hormone) from the hypothalamus, ACTH (adrenocorticotropic hormone) from the pituitary gland and glucocorticoid hormones (in primates, cortisol) from the adrenal cortex.

Rearing — Refers to the practice of raising captive laboratory animals under specific experimental conditions including (1) "mother-reared", or (2) "nursery-reared" in which subjects are permanently removed from their mothers within days of birth, after which the infant may be placed with age-matched peers (peer-rearing), or raised on an inanimate surrogate mother.

Rhesus macaque — Rhesus macaques (*Macaca mulatta*) are a species of Old World Monkey with broad indigenous distribution from eastern Afghanistan to Eastern China. They live in gregarious, hierarchical social groups centered on a core of related lines of females. Rhesus are the most common laboratory primate in the United States.

On the basis of the available evidence from human studies, it is well established that stress early in life, particularly during infancy and early childhood, is a significant predictor and risk factor for a range of developmental, neurological, and behavioral deficits, often including autism-like behaviors. Some deficits manifest early in life, resulting in children who are fearful and anxious, socially withdrawn, or exhibit inappropriate and aggressive play behavior. These adverse experiences and impoverished environments in early human childhood are also known to greatly increase the risk of long-term neurobiological effects on systems essential for normal emotional growth, social competence, and appropriate development of neuroendocrine responsivity. It is clear that experience-based brain development in the very early years of life sets biological pathways that affect cognition, behavior (e.g., anxiety and aggression), capacity to learn, memory, and physical and mental health throughout the life cycle. The hypothalamic-pituitary-adrenal (HPA) axis is exceptionally vulnerable to the effects of early life stress, and such perturbations are suspected of playing a major role in the etiology of affective disorders, mood disorders, and other psychopathologies in adulthood. Delineating the proximal causal links among early experience, brain development, and normative and dysfunctional social behavior is a major goal. Although many questions remain, these biological pathways are becoming increasingly better understood, due largely to evidence that early stressful experience reduces neural plasticity and may lead to the permanent silencing of genes critical to the regulation of the stress response to stress experienced later in life.

Nonhuman Primate Models of Early Stress

Nonhuman primates are a remarkably important model for exploring and ultimately attempting to understand the impact of early stress and trauma on the developing brain, its effects on individual differences in temperament and psychopathology, and the potential interaction of these factors with underlying genetic differences that may impart resilience or vulnerability to early stress. In most primate species, particularly those with gregarious and hierarchical social organization such as rhesus macaques (*Macaca mulatta*), infants are nurtured and protected by their mothers and extended female relatives during their first years. They spend most of their first few months in direct physical contact with their mothers, and even a brief separation can elicit robust behavioral, neuroendocrine, and neurobehavioral activation in both mother and infant.

Studies using animal models such as rodents and nonhuman primates have enabled researchers to address experimentally the effects that early traumatic experiences and impoverished environments can have on a multitude of behavioral and physiological parameters. Early stress and trauma can be defined in a number of ways, but principally they include abuse and neglect of the infant and also may include stress or illness in the mother both pre- and postnatally. In the laboratory, early stress and trauma are modeled most commonly by manipulating the conditions under which infants are reared. The first studies investigating the impact of early adverse experiences were focused on the effects of permanent separation of the infant from its mother – a paradigm initiated by Harry Harlow in the 1960s. Variations of the maternal separation paradigm grew to examine the effects of surrogates on maternally deprived infants and the effects of varying exposure to like-aged peers, but all models emphasized the permanent separation of an infant from its mother.

Although these experiments revealed the devastating consequences of maternal deprivation, they nevertheless represent intense manipulations of early social and emotional experience. As such, they may not appropriately capture the less tangible effects on brain development resulting from the kinds of parental neglect and abuse reported in human clinical populations that increase the risk for adult onset psychopathologies. Alternatively, or in addition to, permanent maternal deprivation are studies investigating the effects of acute, periodic, or repeated mother–infant separations followed by reunion. Compared to the severe social and emotional deficits associated with permanent maternal deprivation, infants that experience short-term maternal separation exhibit more subtle impacts on neuronal and behavioral development. Regardless, long-term physiological and behavioral effects result, including alterations to the HPA axis and in stress reactivity. Indeed, infant monkeys exposed to repeated separations from their mothers show pronounced and persistent behavioral and physiological activation. These experiments in nonhuman primates are more likely to reflect the forms of stressful, unpredictable, and traumatic experiences that neglected and abused children may experience and suggest how such early trauma can affect the normal development of neuronal pathways and neural structures, as well as neuroendocrine responsivity.

Gene by Environment Interaction

The nonhuman primate model is also well suited for studying the interplay between heritable genetic

differences between individuals and their response to nongenetic, environmental insults, that is, gene by environment interaction. Compared to rodents, nonhuman primates such as rhesus macaques are closely related to humans, and this is evident in the degree of similarity between human and rhesus gene sequences of common behavioral candidate genes. At the same time, captive laboratory primates share tightly controlled housing conditions that help to limit the number of intangible, nongenetic variables, which are difficult to isolate in human studies. Finally, most psychopathology involves deficits in appropriate emotional and social functioning, and the complex social organization and behavioral repertoires of nonhuman primates approximate far more closely those of human subjects than do other laboratory animals.

Recent efforts to understand the interaction between genetic and nongenetic (environmental) factors that underlie interindividual differences in response to stress and trauma have focused on functional polymorphisms in genes that are thought to play an important role in mediating response to stress or to be influenced by the effects of early stress and trauma. One of the better-known examples comes from the serotonin system. Serotonin is an important neurotransmitter involved in the regulation of mood and affect, and perturbations in serotonergic neurotransmission, whether naturally occurring or induced, are thought to result in a host of emotional and behavioral disorders in humans. Serotonin is critical for the development of the central nervous system (CNS) during the early postnatal period when mother–infant bonds are cemented. Moreover, serotonergic neurotransmission is involved in the activation and feedback control of the neuroendocrine stress axis, and the activation of stress hormones is known to regulate gene expression, including the serotonin transporter gene (5-HTT).

Both the rhesus macaque and human versions of the serotonin transporter gene contain a complex length polymorphism in the gene's promoter region in which the shorter (s) of the two common alleles reduces transcriptional efficiency. The s allele in humans is associated with numerous mood and psychiatric disorders, including anxiety and depression, harm avoidance, and aggression-related traits. In a recent human longitudinal study, serotonin transporter gene variation increased the risk of depressive symptoms in individuals who had stressful, impoverished childhood experiences. In rhesus macaques, several studies have demonstrated an interaction between the s allele of the serotonin transporter and permanent and acute maternal separation on (1) cerebrospinal fluid measures of 5-hydroxy-3-indoleacetic acid (5-HIAA), a marker of central serotonergic neurotransmission; (2) neonatal measures of motor coordination, reflex, and orienting ability and temperament (thought to be sensitive measures of CNS growth and maturity in the developing brain); and (3) HPA axis activation. In all cases, the effects of serotonin genotype were evident in maternally separated monkeys (but not in maternally raised monkeys), with persistent, lasting effects into adulthood on affective behaviors.

The monoamine oxidase A gene (MAOA), a common candidate gene in human psychiatric genetics involved in metabolizing serotonin, dopamine, and norepinephrine, also contains a functional length polymorphism in the regulatory region that is conserved across humans and rhesus macaques. As with the serotonin transporter gene, the association between high- and low-activity alleles and aggressive behavior was dependent on whether subjects had been raised by their mothers. That is, the effect of functional variation in MAOA was conditioned by early rearing experience.

Although early traumatic and stressful experiences in infancy and childhood have demonstrated effects on a wide range of downstream emotional, behavioral, and neurobiological deficits, it is also clear that not all individuals exposed to early life stress undergo such deficits. The nonhuman primate model for early stress and trauma provides an experimental vehicle for controlled hypothesis-driven analyses that have begun to delineate the complex interactions between early stress and normal gene variation and their effects on brain and CNS function. Yet the specific and causal links between genes and environment, and their interactions on neuronal growth and development in the infant brain, remain poorly understood. As such, there is a continued need for clinical and population studies in humans and in nonhuman animal model systems such as primates to elucidate the complex interplay of heritable predisposition and experience on brain development and function.

See Also the Following Articles

Aggression; Homosexuality, Stress and; Genetic Polymorphisms in Stress Response; Monoamine Oxidase; Serotonin Transporter Genetic Modifications.

Further Reading

Barr, C. S., Newman, T. K., Becker, M. L., et al. (2003). The utility of the non-human primate model for studying gene by environment interactions in behavioral research. *Genes, Brain and Behavior* **2**, 336–340.

Bennett, A. J., Lesch, K. P., Heils, A., et al. (2002). Early experience and serotonin transporter gene variation

interact to influence primate CNS function. *Molecular Psychiatry* **7**, 118–122.
Champoux, M., Bennett, A., Shannon, C., et al. (2002). Serotonin transporter gene polymorphism, differential early rearing and behavior in rhesus monkey neonates. *Molecular Psychiatry* **7**, 1058–1063.
Higley, J. D., Suomi, S. J. and Linnoila, M. (1991). CSF monoamine metabolite concentrations vary according to age, rearing and sex, and are influenced by the stressor of social separation in rhesus monkeys. *Psychopharmacology* **103**, 551–556.
Newman, T. K., Syagailo, Y., Barr, C. S., et al. (2005). Monoamine oxidase A gene promoter variation and rearing experience influence aggressive behavior in rhesus monkeys. *Biological Psychiatry* **57**, 167–172.
Sánchez, M. M., Ladd, C. O. and Plotsky, P. M. (2001). Early adverse experience as a developmental factor for later psychopathology: evidence from rodent and primate models. *Development and Psychopathology* **13**, 419–449.
Suomi, S. J. (1987). Genetic and maternal contributions to individual differences in rhesus monkey biobehavioral development. In: Krasnegor, N. A., Blass, E. M., Hofer, M. A. & Smotherman, W. P. (eds.) *Perinatal development: a psychobiological perspective*, pp. 397–420. New York: Academic Press.

Prison

D L Whitehead and A Steptoe
University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Griffith and A Steptoe, volume 3, pp 241–246, © 2000, Elsevier Inc.

Scope of the Problem
Prison Inmates
Prison Staff
Limitations and Conclusion

Glossary

Bullying	The act of repeatedly intimidating a weaker person by the real or threatened infliction of physical, verbal, or emotional abuse.
Role conflict	A form of social conflict that takes place when the person is forced to take on two different and incompatible roles at the same time.
Depersonalization	A feeling of detachment or estrangement from the self, in which the individual senses a loss of personal identity.
Punitive approach	The perspective on incarceration which reflects the belief that inflicting punishment on an individual will act as a deterrent and decrease the likelihood of the crime being repeated.
Rehabilitative approach	The perspective on incarceration which reflects the belief that an individual can be retrained to function appropriately in society through supportive educational programs in prison.

Scope of the Problem

In 2005, 714/100 000 U.S. residents were incarcerated in prisons and jails, a rate that had risen by 5.0% over the previous 5 years. Bureau of Justice statistics in the United States indicate that if current incarceration rates remain unchanged, an estimated 1 out of every 20 people (5.1%) will serve time in prison during his or her lifetime. U.S. incarceration rates are very much higher than in Europe, where there are wide variations between countries. Southern European countries have a relatively low rate of 80/100 000, compared with 184/100 000 in eastern and central Europe; rates in the United Kingdom fall in the middle, at 139/100 000 in 2005.

There is a very much higher incarceration rate among men than women, and more than 90% of convicts are men. However, one important change over recent years has been the rate of growth of the number of female inmates, which has outstripped that of men. The U.S. female prison population grew by 5.0% annually from 1999 to 2004, compared with a growth rate of 3.3% in the male population. Research has not kept pace with this change, and until recently most studies were conducted on the male inmate population. Overall, inmates are more likely to be young, single, and less well educated than the population at large. Ethnic minority groups are substantially overrepresented in the prison population. In 2004, 12.6% of Black American males in their late twenties were incarcerated, compared with 3.6% of Hispanic males and 1.7% of White males.

Imprisonment is one of the most stressful human experiences and figures high in all comparisons of negative life events. A combination of environmental

and personal deprivation in unpleasant and often frightening circumstances can lead to the perception of extreme stress. Factors such as adherence to a strict daily routine, close living proximity, loss of control over life course, and loss of freedom have all been identified as causing psychological distress. Interpersonal relationships with prison officers and other inmates and the strain on relationships with friends and family are also causes of psychological disturbance. Associated factors such as financial hardship, coping with guilt, and dealing with aggressive fellow inmates can contribute to stress-related problems in prisoners.

Prison staff also report a range of stressors associated with the job. The main stressors are the maintenance of order and the risk of loss of control over hostile inmates, combined with the lack of support from supervisors and society. The difficulty in balancing the responsibility for the welfare of prisoners with the responsibility for security leads to conflicting work demands.

Prison Inmates

Sources of Stress

A number of studies have identified a range of stressors reported among inmates. These fall into several categories.

Personal relationships The separation from partner and family has been cited as one of the main stressors of serving time in prison. For female inmates, the separation from their children is particularly stressful. The majority of women in prison are mothers and many are also heads of households. The separation from their children has been found to be associated with a wide range of feelings such as guilt, anxiety about welfare of the children, and fear of losing close attachments to the children. This may be one reason why women have been found to suffer more emotional distress in prison than men.

The effects of separation are also evident from the finding that married men show higher levels of distress than single men. Being deprived of intimacy and worrying about the welfare of their family are also significant. The lack of confiding relationships and being denied meaningful contact with the opposite gender may create new anxieties regarding sexual identity. Coming to terms with such a mixture of emotions can pose a risk to mental health.

Establishing friendships within prison is an important aspect of adaptation. One study of young inmates found that social support and the ability to confide in friends in prison were major predictors of psychological well-being. The fear of rejection by other inmates and social isolation can cause heightened anxiety, especially in the initial stages of incarceration. However, in many cases such concerns are ill-founded because one study found that prisoners tend to receive more social support from other prisoners than they had anticipated.

Economic factors The loss of income associated with incarceration can be a major problem for prisoners and their families. Men who report financial hardship experience greater distress than those who are financially secure.

Prison environment Lack of privacy and crowding are serious problems within the prison environment. In many countries, prisons are filled beyond capacity. In the United Kingdom, for example, the Home Office released figures in 2005 that indicated a 106% occupancy rate across all prison service establishments. When personal space is threatened, angry or defensive reactions may result. Research has shown that men living in single cells report liking their accommodation more and perceive themselves having greater control than inmates living in multiple occupancy cells or dormitories. Violence is a serious concern for many inmates, and many have long-lasting fears of attack that may be associated with sustained neuroendocrine and autonomic activation. The rates of sexual and physical assault are difficult to estimate because much abuse goes unreported. Bullying and favoritism by staff are also found in many institutions.

Other aspects of the prison environment that can be sources of stress are noise levels, lack of control over lighting, physical discomfort arising from extremes of heat and cold, lack of contact with the natural environment, absence of intellectual stimulation, and constraints on physical activity.

One issue that is problematic is the pressure not to show evidence of distress. Often emotions associated with stress such as fear, grief, and depression are not admitted and are hidden from staff and other inmates. This hinders the analysis and solution of stress-related problems.

Time Course of Imprisonment

Certain sources of stress are more evident at different stages of imprisonment. Three phases of a prisoner's career have been identified.

Initial phase This is the most stressful period when the inmate first encounters prison life. First-time

offenders are particularly vulnerable due to the loss of liberty and separation from friends and family. This stage is often characterized by feelings of denial followed by anxiety and depression, although other emotions such as shock, fear, isolation, grief, and anger are also reported frequently. The adjustment period typically lasts for the first 4–6 weeks of confinement. It is at this stage that many suicide attempts are made.

Middle phase During the middle stage of the sentence, the focus is on the prison community. Friendships are established, and some inmates experiment with drugs and gay sex. Mandatory drug testing has been introduced in many countries; despite efforts to prevent the importation and use of illegal drugs into prisons, the problem is common. In 2001, a UK prison survey found that 19% of inmates reported using cannabis in the previous month, and 13% reported using opiates. Although drug use is prevalent, a smaller proportion of inmates use drugs while in prison than did prior to incarceration.

Final phase Some convicts become preoccupied with their release in the final few weeks of their prison terms. Apprehension about cultural changes after long sentences is coupled with worry about the attitudes of friends and neighbors and the need to reestablish family relationships and reenter employment. Health and economic disparities that existed before incarceration are intensified following release, because access to jobs, health insurance, public housing and health benefits may be limited to ex-prisoners, particularly in the United States, and disproportionately in ethnic minority groups. Thus incarceration continues to exert effects leading to increased stress and poor health even after release.

Responses to the Stress of Prison Life

The stress experienced by prisoners has been measured in several ways: by self-report measures of mood, well-being, and mental health; by the incidence and prevalence of physical health outcomes, mental health outcomes, and physiological indicators of stress; and by the most powerful indicator of hopelessness and despair, suicide.

The investigation of the psychological state of inmates and stress-related effects in prisons faces a number of methodological challenges. First, assessing the impact of incarceration is complicated by limited knowledge about the prior health and well-being of inmates. Many studies of stress in prisons that use outcomes such as hospitalization for mental illness during incarceration are problematic in that incident psychopathology, that is, new onset mental illness, may be difficult to distinguish from prevalent psychopathology, which has been consistently demonstrated as having higher rates in prisoners. Longitudinal studies following prisoners from the first days of incarceration are the only way of studying the emergence of stress resulting from the prison environment rather than ongoing stress from psychosocial disadvantage or existing psychopathology that entails greater risk of conviction in the first place.

Second, the general population is a poor comparison group when assessing the relative prevalence of stress-related disorder in prisoners. Many individuals sent to prison have histories of psychological problems and drug abuse prior to sentencing. Inmates also tend to come from low-income families, to be unemployed, and to have limited education and therefore may have previously encountered negative life events, inadequate health care, and poor living conditions. Drug and alcohol problems frequently precede incarceration; a survey of convicts in state prisons in the United States indicated that 50% of those convicted of violent crimes were under the influence of drugs or alcohol at the time of the offense, and in 1993 drug offenders accounted for 61% of inmates in federal prisons. The high prevalence of some problems among inmates cannot, therefore, be attributed to the prison experience and is confounded by considerable preexisting psychosocial disadvantage in this group.

Nevertheless, anxiety, depression, and complaints of fatigue, sleeplessness, headache, and backache are common. In one study of female convicts, nearly half met the diagnostic criteria for posttraumatic stress disorder (PTSD), and major depression was more prevalent than in the general population. A study of a representative sample of imprisoned juvenile delinquents in California found that half experienced PTSD, with many having witnessed or been involved in traumatic violent episodes. The prevalence of infectious disease, sexually transmitted disease, and tuberculosis is also high compared with the general population. Substantial weight gain has been recorded among female inmates, resulting from the combination of the sedentary lifestyle and the availability of energy-dense high-fat foods. A recent British study investigating subjective health status in female inmates addressed the confounding by lower socioeconomic status in prisoners. Using a comparison group of women from lower-socioeconomic-status sectors of the community, a comparable level of physical health complaints and a greater level of mental health complaints were found in prisoners.

Blood pressure has been found to be elevated in prison inmates compared with matched controls and to increase with the duration of confinement and the degree of crowding in prison. One particularly impressive study used a quasi-experimental approach by measuring the blood pressure of male convicts assigned to individual cells, four-bed cells, and larger dormitories in a medium-security prison. Blood pressure increased to a greater extent in men who were transferred to multiple-occupancy cells and dormitories compared with those who were transferred to other individual cells.

However, a third methodological problem with studies of physiological stress markers is that prison studies have not kept pace with recent advances in measuring such markers as ambulatory blood pressure, cortisol, catecholamines, and inflammation. One notable exception is the Copenhagen Solitary Confinement study, which assessed changes in mood, orientation, and serum cortisol during the first 3 months and compared prisoners kept in solitary confinement (SC) with prisoners in shared cells. Although anxiety, depression, and morning cortisol levels were reduced over time in the non-SC group, such reductions did not occur in the SC group, indicating a lack of adaptation to the prison environment. Many physiological markers of stress are good prognostic indicators of health problems such as autoimmune and cardiovascular diseases, and therefore it may be fruitful to pursue this avenue.

International surveys have found that the suicide rate in prison is between 3 and 11 times greater than that of the general population, with a reported fivefold increase compared with the UK general male population from 1978 to 2003. Most suicides take place within the first few weeks of incarceration and are more common in crowded institutions than other prisons. High rates have been recorded among prisoners on remand and those convicted of murder and serving life sentences. Although suicide is most frequent in young males, self-harm is highly prevalent in female inmates, and recent UK Home Office statistics report a rate of 58.7%. Unlike the pattern in the general population, depression is not a common diagnosis at the time of death, although about one-third of male suicides had some form of psychiatric treatment in the past. Drug or alcohol dependence can be found in a substantial proportion of these cases.

There is also evidence for a European trend toward reinstitutionalization of those with mental health problems in recent years – but in prisons rather than in psychiatric hospitals. A systematic review of prevalence studies in Western countries in 2002 found an overall prevalence of 3.7% for psychotic disorders (10% for major depression), representing a two- to fourfold increase in mental illness compared with the general population. It is likely that U.S. prisons and jails hold up to twice as many individuals with mental illness than do state psychiatric hospitals; the prison system is not well equipped to support these individuals.

Characteristics of the Individual Convict

Some individuals adapt to incarceration better than others and survive even lengthy prison terms without experiencing high levels of distress. Older convicts tend to have more friends in prison, to have more contact with family and friends outside, and to report greater psychological well-being than younger individuals. A recent UK Home Office report, however, drew attention to the often overlooked but growing population of prisoners aged 60 and over. Although physical health problems in this cohort are often well provided for, psychiatric problems, in particular major depression, are not well treated. Political demand for longer sentencing for some crimes means that this older cohort will continue to provide an ongoing challenge for prison services. As noted earlier, higher levels of anxiety and depression are found in female than in male prisoners, although it is not clear whether this is a specific response to imprisonment or reflects the more general difference in distress found in the adult population at large. Members of ethnic and religious minorities suffer more than members of the majority because prejudice and cultural isolation are added to the other sources of stress present in prison.

In the United States, it has been noted that White prisoners are more likely to attempt and complete suicide than African American prisoners, but the reasons are poorly understood. One suggestion is that being raised in the tough conditions experienced by many African Americans makes adults more resilient to the pressures of prison life. Alternatively, because proportionately more Black than White Americans are incarcerated, they may come from a broader spectrum of society; White inmates may include a higher proportion of marginalized, psychologically vulnerable individuals.

Methods of Reducing Stress among Inmates

The management of stress in prison inmates involves several different levels of intervention. Medical and psychiatric services have an important role in preventing suicide and self-harm, in managing drug dependence, and in treating other psychiatric and neurological problems. Support is often hampered by the reluctance of inmates to divulge personal information to authorities or to reveal weakness. Training prison staff more generally in suicide prevention

is also important. A recent public health initiative in Texas prisons led to improvements in health outcomes, and Telemedicine, allowing rapid access to telephone consultations for physical and mental health problems, proved to be a novel and effective way of delivering health care.

At a more general level, education, training, and other initiatives help build prisoners' self-esteem and confidence in their ability to adapt to civilian life after release. Increasing physical activity and exercise may also be beneficial. Studies have found that inmates who have a job detail or are on work-release programs show less evidence of stress than those who are not active.

Finally, changes at the institutional level may be warranted. It is likely that modifications of prison regimes in relation to aspects such as staff attitudes, privileges, training opportunities, and relocation of inmates to serve their terms near their families could play a key role in alleviating stress among inmates. However, the reality is that stress does not figure highly among the concerns of prison administrators or politicians, some of whom would argue that prison should be stressful.

Prison Staff

Sources of Stress

Research on prison staff members finds that many of the issues cited in the general literature on occupational stress are relevant for them as well. As with other jobs, factors such as work demands, role conflict, low control, and poor social support emerge as relevant. Thus, symptoms of psychological distress and job dissatisfaction are greater among correctional officers who report high work demands coupled with low control and low support from management and colleagues.

However, these problems are exacerbated by the peculiar requirements of working in prison. Prison officers cite the conflict of responsibilities in the job as an underlying source of stress and dissatisfaction. On the one hand, prison staff members are responsible for the welfare of inmates and for the maintenance of their physical and psychological well-being. On the other hand, they are required to maintain security and protect the public from the dangerous and sometimes violent convicts. The prison environment is often volatile, with a high risk of aggressive behavior from inmates. Confrontations with inmates and the threat of physical danger are cited as major sources of stress, irrespective of the age and gender of staff members. Ethnic minority prison officers have been found to experience lower stress levels, especially in the United States, where ethnic minorities are grossly overrepresented among inmates.

Some differences have been observed between new and long-serving staff. Newly appointed staff members report a lack of support from supervisors and a concern with lack of control over the working environment as particularly problematic. Those who have been working for longer periods are more troubled by the absence of support and respect from society and the lack of variation in the job. The combination of tedium associated with mundane day-to-day activities and the pressure to maintain vigilance in order to react quickly to high-risk incidents appears to be especially stressful.

Many prison officers believe that the responsibilities for security and the welfare of inmates are incompatible obligations. A recent meta-analysis of factors predicting work-related stress revealed an interesting effect of punitive versus rehabilitative attitudes – officers in the United States who endorsed rehabilitative attitudes experienced greater stress, whereas those in Canada with stronger punitive attitudes were more stressed. This suggests that when individuals' beliefs about their role are at odds with the ethos of the prison service in a particular country, stress and conflict are the result.

Outcomes of Stress

High rates of sick leave for both minor complaints (fatigue, headaches, and indigestion) and physical health problems (high blood pressure and peptic ulcers) occur among prison staff. In one study, the rates of coronary heart disease, hypertension, and ulcers were found to be greater among correctional officers than a comparable sample of police officers. The divorce rate in correctional officers is reported as being twice that of blue- and white-collar workers in general. Job dissatisfaction is coupled with the increased prevalence of divorce in this occupational group. The rates of sick leave are greater among staff members in maximum security prisons and in prisons with many drug users than in other correctional institutions.

There is a greater prevalence of psychological distress in prison officers. Much of the research has focused on burnout, with male and female staff members reporting similar high levels of depersonalization and emotional exhaustion. Self-reported anxiety and depression are also elevated compared with the general working population. Female officers are more likely to take sick leave than male officers, but this pattern may result from family demands rather than different levels of stress at work.

Methods of Reducing Stress in Prison Staff

It has been noted in the research literature that many prison officers do not perceive themselves to be stressed but, rather, see their colleagues as being stressed. Pressure not to show the effects of stress may be particularly evident in a male-dominated environment where there is fear that any sign of weakness will be exploited by the inmates. Admitting to feeling stressed may be interpreted as an inability to cope and therefore as not being strong enough for the job at hand. Alternatively, some prison officers may not be able to identify the personal effects of stress in themselves. This suggests that education about stress and how it can be managed may both benefit prison staff members directly and help them to identify the repercussions of stress in inmates.

A number of studies have found that exercise and fitness programs have positive effects. One investigation involving a 6-week strength and aerobic program showed reductions in body weight, skinfold thickness, and cholesterol levels, together with beneficial changes in smoking, alcohol consumption, sleep, nutritional habits, and stress tolerance.

Consultation and efforts to enhance social support at work have been found to counteract the effects of stress. Factors such as goal consensus among staff, a proactive management style, and increasing of decision latitude can be incorporated into the working environment.

Limitations and Conclusion

There are a number of limitations in the studies that have been carried out thus far about stress in prisons. Most have been cross-sectional in nature, identifying associations between aspects of prison experience and diminished well-being, and the causal sequence is uncertain. Longitudinal designs and data on inmates prior to imprisonment are important in the assessment of the impact of stressors because some apparent effects of incarceration may be the direct consequences of prior experience. The nature of the offense, family situation, and previous physical and mental health are all important factors that should be taken into account. Also, the civilian groups with which prisoners are compared in many studies may not be matched appropriately in terms of background characteristics.

The generalizability of results across populations is uncertain. The majority of studies have been conducted on White male adults and may not be relevant to women, ethnic minorities, and juveniles. This is not mere tokenism because it is apparent from the limited information available that the experience of prison is very different for people of different backgrounds and cultures. The applicability of findings to different prison systems must also be questioned. Justice systems in the United Kingdom and other European countries differ greatly from the U.S. system, and countries also vary in the extent to which individuals with serious psychiatric histories are found in prison. The application of management suggestions derived from one system must be reviewed carefully before their implementation elsewhere.

Nevertheless, the data do indicate that imprisonment is severely stressful for many inmates and that work within the prison system is also associated with higher rates of stress-related dysfunction than comparable jobs. Education and the implementation of specific stress-management programs may be beneficial to both groups and may help ensure that stress in prison is confined within tolerable limits.

See Also the Following Articles

Burnout; Crowding Stress; Prisoners of War; Racial Harassment/Discrimination; Workplace Stress.

Further Reading

Andersen, H. S. (2004). Mental health in prison populations: a review. *Acta Psychiatrica Scandinavica* **110**(supplement 424), 5–59.

Andersen, H. S., Sestoft, D., Lillebaek, T., et al. (2003). A longitudinal study of prisoners on remand: repeated measures of psychopathology in the initial phase of solitary versus nonsolitary confinement. *International Journal of Law and Psychiatry* **26**, 165–177.

Cooke, D. J., Baldwin, P. J. and Howison, J. (1993). *Psychology in prisons*. London: Routledge.

DeRosia, V. R. (1998). *Living inside prison walls: adjustment behavior*. Westport, CT: Praeger.

Hayes, L. M. and Blaauw, E. (eds.) (1997). *Special issue on prison suicide. Crisis: The Journal of Crisis Intervention and Suicide Prevention* **18**(4).

Dowden, C. and Tellier, C. (2004). Predicting work-related stress in correctional officers: a meta-analysis. *Journal of Criminal Justice* **32**, 31–47.

Lindquist, C. H. and Linquist, C. A. (1997). Gender differences in distress: mental health consequences of environmental stress among jail inmates. *Behavioral Sciences & the Law* **15**, 503–523.

Ostfeld, A. M., Kasl, S. V., D'Atri, D. A. and Fitzgerald, E. F. (1987). *Stress, crowding and blood pressure in prison*. Hillsdale, NJ: Lawrence Erlbaum.

Paulus, P. B. (1988). *Prison crowding: a psychological perspective*. New York: Springer-Verlag.

Prisoners of War

C Tennant
University of Sydney, St. Leonards, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 247–251, © 2000, Elsevier Inc.

Mortality, Physical Morbidity, and Cognitive Impairment
Psychological Morbidity
Etiological Factors
Summary

Glossary

Physical morbidity	Illness and hospital admissions.
Psychological morbidity	Anxiety, depression, posttraumatic stress disorder, and psychiatric admissions.

Mortality, Physical Morbidity, and Cognitive Impairment

Mortality

There have been relatively few studies of physical health and mortality in prisoners of war (POWs) and controls; studies from different countries have produced different findings. The largest series of studies is based on a cohort of U.S. World War II Pacific, European, and Korean POWs and controls assessed by a series of researchers. In the early postwar years, there was increased mortality in all POW groups, due largely to tuberculosis (TB) and postwar trauma. By 1965, no increased mortality appeared in European POWs. Mortality was still increased in Pacific POWs, due to cirrhosis, tuberculosis, and trauma, and in Korean veterans, due to trauma. This increased mortality in Pacific and Korean POWs had persisted for 9 and 13 years, respectively, postwar, whereas death from cirrhosis of the liver had increased from the tenth year of follow-up. In an Australian study of POWs of the Japanese in the 1950s, more deaths from suicide, trauma, cirrhosis, and TB were initially reported. In an Australian controlled cohort study (40 years postwar), no increase in mortality in POWs was found.

Physical Health and Hospital Admissions

In U.S. POWs from all three theaters, general hospital admission rates to 1965 were increased significantly. However, in a 40-year follow-up of a subsample, no increased rates of hypertension, diabetes, or myocardial infarction were found. A full medical examination of the Australian cohort revealed no differences in current medical disorders. Higher rates of peptic ulcer were, however, generally found in World War II POWs and in U.S. naval POWs in Vietnam; the latter also experienced excessive disorders of the peripheral nervous system.

Thus, postwar peptic ulcer, trauma, TB, and cirrhosis appear more common in World War II POWs than in combatant controls. Although mortality appears to be increased in the U.S. studies, there may have been sampling biases; findings from a well-matched Australian cohort study suggested no difference.

Cognitive Impairment

U.S. World War II and Korean prisoners of war High- and low-weight-loss POWs have been compared with combatant controls. Cognitive impairment (memory and learning) was greatest in POWs with the greater weight loss; there was, however, little difference between low-weight-loss POWs and their controls.

Australian prisoners of war of the Japanese No significant differences in the Australian cohort follow-up (50 years postwar) were found between POWs and combatant controls using detailed psychometric tests; high- and low-weight-loss groups similarly did not differ on any tests or on prevalence of clinically diagnosed dementia. These similarities occurred despite POWs having a greater incidence of depression that could have impaired their psychometric performance, a factor that may have accounted for the positive findings in U.S. studies, which relied on volunteers. Finally, in the same cohort, POWs and controls did not differ on any neuropsychological tests or on brain width (using cerebral X-ray computed tomography, CT, scan); however, POWs, interestingly, had a shorter mean anteroposterior brain diameter. Taking into account methodological problems of some studies, cognitive impairment is not confirmed in POWs postwar.

Psychological Morbidity

Minnesota Multiphasic Personality Inventory Questionnaire Studies (U.S. World War II and Korean Prisoners of War)

U.S. World War II prisoners of war Minnesota Multiphasic Personality Inventory (MMPI) symptom profiles (depression, anxiety, and somatic concern)

were elevated (from highest to lowest) in POWs with chronic posttraumatic stress disorder (PTSD), POWs recovered from PTSD, POWs with other psychiatric diagnoses, POWs with no psychiatric disorder, and normal men. The personality profiles of denial and suppression were also elevated.

Korean prisoners of war and combatant controls Hypochondriasis, depression, hysteria, psychasthenia, paranoia, and ego strengths distinguished the POWs and the controls. The MMPI assesses both personality and psychological symptoms. It shows some unexpected personality differences in POWs and controls, presumably because the personality subscales are sensitive to change in mood. The Eysenck Personality Questionnaire (a more stable measure) showed no differences when used in other Australian POWs and controls.

Psychiatric Admissions and General Psychological Morbidity Data

U.S. prisoners of war (World War II and Korea) Hospital admission rates were significantly higher in the Pacific POWs for 27 of the 33 somatic conditions studied and for all nine of the psychiatric diagnoses, whereas in Korean POWs, admissions were elevated in only nine somatic conditions and three psychiatric conditions. Admission rates in the European POWs were not increased. Bias was possible due to a response rate of only 59% for the initial medical questionnaire. Again, 45 years postwar these POWs had moderately elevated rates of hospitalization and depressive disorders and greatly elevated PTSD; no difference was found in bipolar disorder, schizophrenia, alcoholism, or physical health.

Australian World War II prisoners of war of the Japanese These POWs (88% response) were no more likely to have had psychiatric admissions than combatant controls and indeed had fewer postwar multiple psychiatric admissions. Diagnosed clinical anxiety and major depression in POWs were twice that of controls; alcohol abuse and psychosis were no different. These findings were consistent with their significantly higher scores on questionnaire measures of anxiety and depression. Surprisingly, 10 years later there had been a significant decline in both diagnosed clinical anxiety and depression.

General psychological morbidity is elevated significantly for many years postwar; however, admission data findings vary between U.S. and Australian POWs.

Posttraumatic Stress Disorder

World War II and Korea

U.S. prisoners of war This group has been studied extensively by Engdahl and colleagues and by Sutker and colleagues. Initial questionnaires were returned by only 35% of the sample surveyed; some 500 of these were then studied 40 years postwar with a 75% response. In the first series of 426 POWs, the rate of PTSD was increased greatly and lifetime prevalence of depression was increased moderately, but bipolar disorder, schizophrenia, and alcoholism were not. In 262 of this sample, 53% met the criteria for lifetime PTSD and 29% for current PTSD; in Japanese POWs, the rates of PTSD were higher, 84% lifetime and 59% current PTSD prevalence. In 62 of the World War II POWs, 50% had PTSD within 1 year of release and 29% has PTSD 40 years later. A lifetime diagnosis of generalized anxiety was found in over 50%. In another study of 915 of the original World War II cohort, 53% of POWs initially responded and 188 of these were interviewed. Sixty-seven percent of this sample had PTSD; 50% had shown no recovery, 24% had some residual symptoms, 39% had mild symptoms, and 29% had fully recovered. It is important to note the possibility of considerable sample selection bias in this series of studies (due to poor response rates), which may have inflated the rates of psychiatric morbidity.

The second U.S. series of studies are those of Sutker and colleagues. Korean and Pacific veterans ($n = 326$) had experienced more extreme trauma and had the highest prevalence rates for lifetime and current mental disorders, including PTSD. In a small subsample, 77% of the Pacific POWs had current PTSD and 78% a lifetime diagnosis, compared with 18 and 29%, respectively, in controls. In Korean POWs, PTSD symptoms were found in 90–100%. Finally, World War II aviators were less likely to have PTSD than nonaviators. The sample bias was also probable in these studies because the subjects were selected from attenders at a Veteran's Administration (VA) medical center. Other studies have used questionnaires to assess PTSD with similar high prevalence rates in military and civilian POWs.

Canadian prisoners of war Higher rates of PTSD and other symptoms were found in POWs from Europe 50 years postwar, whereas in former Far East POWs, 30% had current PTSD 50 years postwar and 90% had at least one PTSD symptom.

Australian prisoners of war In the Australian cohort, no subjects appeared to meet full *Diagnostic and Statistical Manual of Mental Disorders,* 3rd ed. (DSM-III) criteria for PTSD 40 years postwar, although a majority of the sample had some PTSD symptoms and 50% had a clinical affective disorder or anxiety. At a further follow-up 10 years later, 4% of subjects had been diagnosed as having PTSD using DSM-IV criteria; there was, however,

a significant decline in psychiatric disorders overall, from 50 to 36%.

French and German prisoners of war of the Russians Forty years postwar, 71% of a selected sample (response 40%) had symptoms consistent with PTSD, whereas in a randomly selected subsample, the frequency of individual PTSD symptoms ranged from 84% (for recurring dreams) to 18% (for panic symptoms).

Dutch Resistance Veterans Sample veterans were studied 40 years postwar (76% response); half had been imprisoned and half had not. Interestingly, 54% of both groups had PTSD and a similar rate of other psychological symptoms. Paradoxically, this study followed an earlier study showing that, 15 years postwar, imprisoned Dutch resistance fighters had less incidence of psychiatric disorders than those not imprisoned.

There is little doubt that significant rates of PTSD are found long after the war, although not in all samples; the very high rates of PTSD found, particularly in U.S. samples, may be due in part to various forms of sampling bias.

Other Conflicts

Israeli prisoners of war of the Yom Kippur War POWs, combatants (who had experienced combat stress reactions), and normal combat veterans were compared; the prevalence of PTSD more than 18 years postwar in the three groups was 23, 37, and 14%, respectively, whereas current PTSD prevalence was 13, 13, and 3%. Psychophysiological complaints were also more frequent among POWs, as were trauma-related and general psychiatric symptoms. In terms of coping, those with high sensation-seeking qualities adjusted better to the stress of captivity and had fewer PTSD or other psychiatric symptoms; active coping strategies were used by high sensation seekers, whereas low sensation seekers used detachment and denial. In general, poor psychological adjustment was associated with more helplessness, more suffering, and less active coping, as well as hostility and ambivalence.

British prisoners of war from the Persian Gulf War In POWs studied both 6 and 18 months postwar, psychological symptoms (using the General Health Questionnaire, GHQ) reduced significantly; a poorer psychological outcome was associated with POWs witnessing physical violence and self-perceived deterioration in physical and mental health during captivity.

Balkan prisoners of war In Croatian POWs, 36% had postwar psychiatric disorder; camp conditions contributed to psychiatric symptoms (particularly threats against family members); however, group cohesion diminished the level of disorder. In another sample, only 4.5% met the full criteria for PTSD, 28% had no PTSD symptom cluster, and the remainder had some PTSD symptoms.

Comorbidity of Psychiatric Disorders in Prisoners of War

In one selected sample of Korean POWs, PTSD was found the most; depressive comorbidity occurred in 75%, anxiety in 45%, and alcohol abuse in 20%.

In the other U.S. series, of 262 World War II and Korean POWs (who had sought mental health treatment), 53% met the criteria for lifetime PTSD; 29% of those with PTSD had other axis I disorders. In a subsample, depression and anxiety coexisted in 61% of the sample. Depression was found in 10% of those without PTSD, in 23% of those with recovered PTSD, and in 61% of those with ongoing PTSD. The Australian POW study also found that anxiety and depression coexisted in a significant portion of the sample.

Etiological Factors

The severity of the stressful experiences of POWs differed in different theaters. Both Korean and Vietnam POWs were interrogated, brain-washed, and tortured, whereas Vietnam POWs were also kept for long periods in solitary confinement. POWs of the Japanese suffered gross physical deprivations, malnutrition, random brutality, hard physical labor, major physical illness, and ultimately a high mortality rate. European theater POWs did not suffer as severely, although resistance fighters (such as the Dutch) had a similar experience to Holocaust survivors. The POW experience in more recent conflicts has generally been briefer and less brutal. Studies addressing etiology have been somewhat piecemeal; World War II POWs provide the most consistent findings.

World War II Prisoners of War

In U.S. POWs, studied by Engdahl and colleagues, generalized anxiety and depressive disorders were related to captivity severity; preexisting psychopathology and family history of psychiatric disorder were not predictive. Chronic depression was predicted by a younger age at time of capture, lower education, more medical symptoms during captivity, and poorer social supports after release. Specific

captivity stressors (weight loss, torture, and disease during captivity) correlated strongly with psychological morbidity and PTSD, whereas age and education were protective. These findings have been largely replicated in the other U.S. series of studies by Sutker and in Canadian, French, and Australian POWs.

Vietnam Prisoners of War

Captivity trauma was greater in POWs captured before 1969 and was associated with more subsequent psychiatric morbidity. Aviators were more resilient to captivity than nonaviators; this was attributed to their higher educational status and military rank and better personal resources. Furthermore, there was no particular coping style indicative of psychiatric illness in this group. Paradoxically, still-serving POWs reported that the harshness of the war experience was associated with a belief of having actually benefited from imprisonment.

The severity of the captivity experience, especially as manifest in gross physical hardship and deprivation, predicted the subsequent short- and long-term psychological morbidity.

Summary

The POW experience varied according to the theater of war; the Japanese were particularly brutal captors. In the early postwar years, excess physical morbidity and mortality occurred particularly in Pacific POWs. Although this declined subsequently, deaths from cirrhosis showed a delayed increase.

Significant psychiatric morbidity has been the most consistent finding in all the studies, including conflicts since World War II, and morbidity has persisted for many years postwar. Studies in the United States have focused specifically on PTSD and show very high rates of prevalence, but sampling bias may, to some degree, contribute to their particularly high rates. Other anxiety disorders and depression are also increased, but psychoses are not. The comorbidity of disorders is common. Furthermore, the high rates of psychological disorder correlate with a range of indices of captivity severity, whereas some sociodemographic factors seem protective. Former POWs who served their countries well and suffered terribly continue to live with the legacy of the inhumanity of their captors during war.

See Also the Following Articles

Combat Stress Reaction; Concentration Camp Survivors; Korean Conflict, Stress Effects of; Persian Gulf War, Stress Effects of; Prison; Vietnam Veterans, Postwar Experiences and Health Outcomes; War-Related Posttraumatic Stress Disorder, Treatment of.

Further Reading

Beebe, G. W. (1975). *American Journal of Epidemiology* **101**, 400–422.

Bisson, J. I., et al. (1998). *British Journal of Medicine and Psychology* **71**, 247–252.

Eberly, R. E. and Engdahl, B. E. (1991). *Hospital and Community Psychiatry* **42**, 807–813.

Engdahl, B. and Dikel, T. N. (1997). *American Journal of Psychiatry* **154**, 1576–1581.

Engdahl, B. E., et al. (1991). *Journal of Nervous and Mental Disease* **179**, 181–187.

Engdahl, B. E., et al. (1991). *Social, Psychiatry and Psychiatric Epidemiology* **26**, 63–67.

Goulston, K. J. (1985). *Medical Journal of Australia* **143**, 6–10.

Jorm, A. F., et al. (1997). *Personality and Individual Differences* **23**, 371–377.

Kluznic, J. C., et al. (1986). *American Journal of Psychiatry* **143**, 1443–1446.

Nefzer (1970). *American Journal of Epidemiology* **91**, 123–138.

Neria, Y. and Solomon, Z. (1998). *Journal of Nervous and Mental Disease* **186**, 174–182.

Solomon, Z., et al. (1998). *European Journal of Personality* **12**, 271–285.

Sulway, M. R., et al. (1996). *Neurology* **46**, 650–655.

Sutker, P. B. and Allain, A. N. (1991). *Psychological Reports* **68**, 279–284.

Sutker, P. B. and Galina, Z. H. (1990). *Journal of Consulting and Clinical Psychology* **58**, 323–328.

Tennant, C. and Fairley, M. J. (1997). *Journal of Nervous and Mental Disease* **185**, 686–689.

Tennant, C. and Goulston, K. (1986). *American Journal of Psychiatry* **143**, 618–621.

Tennant, C., et al. (1986). *Psychology and Medicine* **16**, 833–839.

Wilson, J. P. and Raphael, B. (eds.) (1993). *International handbook of traumatic stress syndromes*. New York: Plenum Press.

Problem-Solving Skills Training

A M Nezu and C M Nezu
Drexel University, Philadelphia, PA, USA
T J D'Zurilla
State University of New York, Stony Brook, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A M Nezu, C M Nezu and T J D'Zurilla, volume 3, pp 252–256, © 2000, Elsevier Inc.

Social Problem Solving
The Problem-Solving Process
Problem-Solving Deficits
Problem Solving as a Moderator of Stress
Efficacy of PST
PST Training Guidelines

Glossary

Effective solutions	Coping responses that achieve one's problem-solving goals, but that also simultaneously maximize positive consequences (i.e., benefits) and minimize negative consequences (i.e., costs).
Problems	Specific existing or anticipated situations that demand responses for adaptive functioning, but that are not met by effective coping responses from the person confronted by them due to the presence of certain obstacles (e.g., ambiguity, uncertainty, conflicting demands, lack of resources, or uniqueness).
Problem orientation	The set of relatively stable cognitive-affective schemas that represent a person's generalized beliefs, attitudes, and emotional reactions about problems in living and one's ability to successfully cope with such problems.
Problem-solving style	The core cognitive-behavioral activities that people engage in when attempting to cope with problems in living.
Social problem solving	The cognitive-behavioral process by which individuals view the nature of problems in living and direct their attempts at coping with such stressors.
Solution	A coping response designed to alter the nature of the problematic situation, one's negative emotional reactions to it, or both.

Social Problem Solving

Throughout history, psychologists and philosophers have argued that an essential part of being human is the capacity to solve problems. Perhaps the most important idea emanating from this belief is the notion that problem-solving ability contributes significantly to social competence and overall psychological well-being, as the ability to cope with and resolve everyday stressful problems has been found to be strongly related to overall personal and social functioning.

Training individuals to become better problem solvers in order to facilitate their ability to cope with stressful situations has been referred to in the psychotherapy and counseling literature as social problem-solving therapy in order to highlight the social and interpersonal context in which real-life problem solving occurs. Social problem solving is defined as the cognitive-behavioral process by which individuals view the nature of problems in living and direct their attempts at coping with such stressors. Relevant problem-solving therapy (PST) goals broadly include altering the stressful nature of the situation itself (e.g., overcoming obstacles to a personal goal), changing one's reactions of distress to such problems (e.g., acceptance that a goal cannot be reached), or both.

Problems can be single events (e.g., obtaining a loan from a bank), a series of related problems (e.g., continuous arguments with a spouse), or chronic situations (e.g., a major chronic illness, such as cancer). Situations become problems when effective responses are required in order for the person to cope adaptively, but when such responses are not immediately available or identifiable due to the presence of various obstacles, including ambiguity, unpredictability, conflicting demands, deficient skills, or lack of resources.

Solutions are coping responses designed to alter the nature of the problematic situation, one's negative emotional reactions to it, or both, whereas effective solutions are those coping responses that not only achieve such goals, but also simultaneously maximize positive consequences (i.e., benefits) and minimize negative effects (i.e., costs).

The Problem-Solving Process

Problem-solving outcomes are largely determined by two general but partially independent dimensions – problem orientation and problem-solving style. Problem orientation is the set of relatively stable

cognitive-affective schemas that represent a person's generalized beliefs, attitudes, and emotional reactions about problems in living and one's ability to successfully cope with such problems. Problem orientation can be either positive or negative. A positive orientation is one that involves a tendency to appraise problems as challenges, be optimistic in believing that problems are solvable, perceive one's own ability to solve problems as strong, and believe that successful problem solving involves time and effort. Conversely, a negative problem orientation is one that involves the tendency to view problems as threats, expect problems to be unsolvable, doubt one's own ability to solve problems successfully, and become frustrated and upset when actually faced with problems.

Problem-solving style refers to those core cognitive-behavioral activities that people engage in when attempting to cope with problems in living. Three differing styles have been identified, one that is adaptive, while the other two reflect maladaptive ways of coping. Rational problem solving is the constructive problem-solving style that involves the systematic and planful application of certain specific skills, each of which makes a distinct contribution toward the discovery of an adaptive solution or coping response in a problem-solving situation. Rational problem solving involves the following four skills: problem definition and formulation, generation of alternatives, decision making, and solution implementation and verification. The goal of problem definition and formulation is to delineate the reasons why a given situation is a problem (e.g., the presence of obstacles), as well as to specify a set of realistic goals and objectives to help guide further problem-solving efforts. The purpose of the generation of alternatives task is to create, using various brainstorming principles, a large pool of possible solutions in order to increase the likelihood that the most effective ideas will be ultimately identified. The goal of decision making is to conduct a systematic cost–benefit analysis of each alternative by identifying and then weighing their potential positive and negative consequences if carried out, and then, based on this evaluation, to develop an overall solution plan. Finally, the purpose of solution implementation and verification is to carry out the solution plan, monitor and evaluate its effectiveness, and troubleshoot if the outcome is unsatisfactory.

Two additional problem-solving styles have been identified, both of which are dysfunctional or maladaptive in nature. An impulsive/careless style involves a generalized response pattern characterized by impulsive, hurried, and careless attempts at problem resolution. Although the individual characterized by this style actively attempts to apply various strategies to address problems, such attempts are narrow, hurried, or incomplete. For example, a person with this style is likely to consider only a few solution alternatives, often impulsively implementing the first idea that comes to mind. In addition, the narrow range of options and their consequences are scanned quickly, carelessly, and unsystematically.

Avoidance style is a second maladaptive problem-solving pattern, this one characterized by procrastination, passivity, and overdependence on others to provide solutions. This type of individual generally avoids problems rather than confronting them head on, puts off addressing problems for as long as possible, waits for problems to resolve themselves, and attempts to shift the responsibility for solving his or her problems to other people. In general, both styles lead to ineffective or unsuccessful problem resolution.

Problem-Solving Deficits

Important differences have been identified in individuals characterized as effective versus ineffective problem solvers. In general, when compared to their effective counterparts, ineffective problem solvers report a greater number of life problems, more health and physical symptoms, more anxiety, more depression, and more psychological maladjustment. In addition, a negative problem orientation has been found to be associated with negative moods under routine and stressful conditions in general, as well as pessimism, negative emotional experiences, and clinical depression. Persons with a negative orientation also tend to worry and complain more about their health.

In addition, problem-solving deficits have been found to be significantly related to poor self-esteem, hopelessness, suicidal risk, self-injury, anger proneness, increased alcohol intake and substance risk taking, personalities difficulties, criminal behavior, alcoholism, secondary physical complications among persons with spinal cord injuries, premenstrual and menstrual pain, physical health problems, diminished life satisfaction, physical problems among adult cancer patients, and pain severity among adult cardiac patients.

Problem Solving as a Moderator of Stress

How people cope with stressful experiences, including major events (e.g., undergoing a divorce, dealing with the death of a spouse) and daily problems (e.g., continued arguments with a co-worker, limited financial resources) can, in part, determine the degree to which they will experience long-lasting psychological distress, particularly depression. Continued successful attempts at problem resolution will lead to a reduction or minimization of immediate emotional

distress and a reduced likelihood of long-term negative affect (i.e., clinical depression). Alternatively, if one's problem-solving coping skills are ineffective, or if extreme emotional distress impacts negatively on one's coping efforts, resulting in reduced motivation, inhibition of problem-solving performance, or both, then the likelihood of long-term emotional distress will be increased. Further, such negative outcomes can lead to the exacerbation of existing problems and the creation of new ones, which in turn can lead to another major stressful life event, and so forth. As such, how one copes with problems can lead to either an escalation or attenuation of the stress process. For example, research has demonstrated that under similar levels of high stress, individuals with poor problem-solving skills experience significantly higher levels of depression and anxiety than persons characterized by more effective problem-solving skills, supporting the notion that problem solving serves to attenuate the negative effects of stress.

Efficacy of PST

If effective problem-solving skills serve as an important buffering factor regarding the stress process, training individuals in such skills should lead to a decrease in emotional distress and improvement in overall psychological functioning. In fact, PST has been shown to be effective in a wide range of clinical populations, psychological problems, and the distress associated with chronic medical disorders. These include unipolar depression, geriatric depression, distressed primary care patients, social phobia, agoraphobia, obesity, coronary heart disease, adult cancer patients, schizophrenia, mentally retarded adults with concomitant psychiatric problems, HIV risk behaviors, drug abuse, suicide, childhood aggression, and conduct disorder.

In addition to its applicability to a variety of patient populations, PST also appears to be flexible with regard to treatment goals and methods of implementation. For example, it can be conducted in a group format, on an individual and couples basis, and as part of a larger psychosocial intervention package, and can be delivered by telephone. It can also be applied as a means of helping patients to overcome barriers associated with successful adherence to other medical or psychosocial treatment protocols.

PST Training Guidelines

Specific PST therapy objectives include (1) enhancing individuals' positive orientation and application of the four rational problem-solving tasks, and (2) minimizing their negative orientation and tendency to engage in dysfunctional problem-solving style activities (i.e., impulsive or careless attempts to cope with problems; avoidance of problems). PST interventions include didactic explanations, training exercises, practice opportunities, and homework assignments geared to foster practice between training sessions.

Problem Orientation

Training in this problem-solving component is geared to facilitate the following: positive self-efficacy beliefs (the perception that people can improve their quality of life through effective coping and problem solving), beliefs that problems are inevitable (accepting the notion that it is common and normal to experience a wide range of problems), the ability to identify problems accurately when they occur, and the ability to inhibit emotional reactions that can lead to impulsive reactions or avoidance.

A variety of training approaches can be used to foster a positive problem orientation. One technique is the reverse advocacy role-play strategy. According to this strategy, the therapist pretends to adopt a particular belief about problems and asks the patient to provide reasons why that belief is irrational, illogical, incorrect, or maladaptive. Such beliefs might include the following statements: "Problems are not common to everyone; if I have a problem, that means I'm crazy," "There must be a perfect solution to this problem," "I'll never be the same again." At times when the patient has difficulty generating arguments against the therapist's position, the counselor then adopts a more extreme form of the belief, such as "No matter how long it takes, I will continue to try and find the perfect solution to my problem." This procedure is intended to help patients identify alternative ways of thinking and then to dispute or contradict previously held negative beliefs with more adaptive perspectives.

Patients are also taught to use feelings or emotions as cues that a problem exists by using visual images of a red traffic stop sign as a signal to stop and think. In essence, patients are taught to recognize various situations as problems and to label them as such. Accurately labeling a problem as a problem serves to inhibit the tendency to act impulsively or automatically in reaction to such situations. It also facilitates the tendency to approach or confront problems, rather than to avoid them.

Problem Definition and Formulation

Problem definition can be likened to mapping a guide for the remainder of the problem-solving process. The major focus of this task is to better understand the nature of the problem and to set clearly defined and

reasonable goals. In other words, locating a specific destination on a map makes it easier to find the best route to get there. Training in problem definition and formulation focuses on the following five specific tasks: gathering all available information about the problem, using clear and unambiguous language, separating facts from assumptions, identifying the factors that make the situation a problem, and setting realistic problem-solving goals.

Generation of Alternatives

When generating alternative solutions to a problem, PST encourages broad-based, creative, and flexible thinking. In essence, patients are taught various brainstorming strategies (e.g., the more the better, defer judgment of ideas until a comprehensive list is created). This helps to increase the likelihood that the best or most effective solution ideas will be discovered.

Decision Making

Once a list of alternative options has been generated, the problem solver begins to systematically and thoroughly evaluate the potential for each solution to meet the defined goal(s). Training in this component helps the patient to use the following criteria to conduct a cost–benefit analysis based on the utility of each alternative solution: the likelihood that the solution will meet the defined goal, the likelihood that the person responsible for solving the problem can actually carry out the solution plan optimally, personal (i.e., effects on oneself) and social (i.e., effects on others) consequences, and short- and long-term effects.

Solution Implementation and Verification

This last rational problem-solving task involves first carrying out the solution plan and then monitoring and evaluating the consequences of the actual outcome. PST encourages the patient to practice the performance aspect of solution implementation as a means of enhancing the probability that it will be carried out in its optimal form. Once the plan is under way, the patient is encouraged to monitor the actual results. Using this information allows the individual to evaluate the results by comparing the actual outcome with his or her expectations or predictions about the outcome.

Supervised Practice

After the majority of training has occurred, the remainder of PST should be devoted to practicing the newly acquired skills and applying them to a variety of stressful problems. Beyond actually solving stressful problems, ongoing in-session practice serves three additional purposes: the patient can receive professional feedback from the therapist, increased facility with the overall PST model can decrease the amount of time and effort necessary to apply the various problem-solving tasks with each new problem, and practice fosters maintenance and generalization of the skills.

The number of practice sessions required after formal PST training often is dependent upon the competency level a patient achieves, as well as on the actual improvement in his or her overall quality of life. In the research protocols that have found PST to be an effective cognitive-behavior therapy intervention, the number of included sessions has ranged from 8 to 12.

See Also the Following Articles

Anxiety; Cognitive Behavioral Therapy; Depression and Manic-Depressive Illness.

Further Reading

Allen, S. M., Shah, A. C., Nezu, A. M., et al. (2002). A problem-solving approach to stress reduction among younger women with breast carcinoma: a randomized controlled trial. *Cancer* **94**, 3089–3100.

Chang, E. C., D'Zurilla, T. J. and Sanna, L. J. (eds.) (2004). *Social problem solving: theory, research, and training*. Washington, D.C.: American Psychological Association.

D'Zurilla, T. J. and Nezu, A. M. (1999). *Problem-solving therapy: a social competence approach to clinical intervention* (2nd edn.). New York: Springer.

D'Zurilla, T. J., Nezu, A. M. and Maydeu-Olivares (2002). *Social Problem-Solving Inventory-Revised (SPSI-R): technical manual*. North Tonawanda, NY: Multi-Health Systems.

D'Zurilla, T. J., Nezu, A. M. and Maydeu-Olivares, A. (2004). Social problem solving: theory and assessment. In: Chang, E. C., D'Zurilla, T. J. & Sanna, L. J. (eds.) *Social problem solving: theory, research, and training*, pp. 11–27. Washington, D.C.: American Psychological Association.

Nezu, A. M. (1987). A problem-solving formulation of depression: a literature review and proposal of a pluralistic model. *Clinical Psychology Review* **7**, 122–144.

Nezu, A. M. (2004). Problem solving and behavior therapy revisited. *Behavior Therapy* **35**, 1–33.

Nezu, A. M. and D'Zurilla, T. J. (1989). Social problem solving and negative affective states. In: Kendall, P. C. & Watson, D. (eds.) *Anxiety and depression: distinctive and overlapping features*, pp. 285–315. New York: Academic Press.

Nezu, A. M. and Perri, M. G. (1989). Problem-solving therapy for unipolar depression: an initial dismantling investigation. *Journal of Consulting and Clinical Psychology* **57**, 408–413.

Nezu, A. M., Nezu, C. M. and Perri, M. G. (1989). *Problem-solving therapy for depression: theory, research, and clinical guidelines*. New York: Wiley.

Nezu, A. M., Nezu, C. M., Friedman, S. H., Faddis, S. and Houts, P. S. (1998). *Helping cancer patients cope: a problem-solving approach*. Washington, D.C.: American Psychological Association.

Nezu, A. M., Nezu, C. M., Felgoise, S. H., McClure, K. S. and Houts, P. S. (2003). Project genesis: assessing the efficacy of problem-solving therapy for distressed adult cancer patients. *Journal of Consulting and Clinical Psychology* **71**, 1036–1048.

Nezu, C. M., D'Zurilla, T. J. and Nezu, A. M. (2005). Problem-solving therapy: theory, practice, and application to sex offenders. In: McMurran, M. & McGuire, J. (eds.) *Social problem solving and offenders: evidence, evaluation and evolution*, pp. 103–123. Chichester, UK: Wiley.

Perri, M. G., Nezu, A. M., McKelvey, W. F., et al. (2001). Relapse prevention training and problem-solving therapy in the long-term management of obesity. *Journal of Consulting and Clinical Psychology* **69**, 722–726.

Prolactin and Stress

G Tolis
Hippokrateion Hospital, Athens, Greece, and McGill University, Montreal, Canada
G Rombopoulos, D Kaltsas, E Katounda, V Kaltzidou and N Angelopoulos
Hippokrateion Hospital, Athens, Greece

© 2007 Elsevier Inc. All rights reserved.

Glossary

Catabolism	A wasting process.
Dopamine	A catecholamine neurotransmitter that mediates nerve signaling in the central nervous system.
Opiergic compounds	Opiates or compounds that have an opiate-like effect, for example, morphine, methadone, and heroin.
Prolactin (PRL)	Polypeptide produced in the pituitary with somatolactotropic actions, essential for the preparation of the breast during pregnancy and to assure postpartum lactation.
Somatomammotropic effect	The growth of breast function regulatory effect.

Prolactin (PRL), a somatolactotropic pituitary hormone of paramount importance in lactation, is under tonic inhibition primarily by the tubero-infundibular dopaminergic neurons (TIDA). Psychotropic medications (the class of butyrophenones, i.e., haloperidol) increase PRL levels, whereas dopamine agonists (e.g., apomorphine and bromocriptine) suppress PRL secretion. Recent advances in neurochemistry have made drugs available that have antipsychotic actions without altering PRL secretion (e.g., clozapine). Thyrotropin-releasing hormone (TRH) increases serum PRL levels by a direct effect on the anterior pituitary gland, as do many of the antipsychotic drugs. The latter act as dopamine receptor antagonists at the pituitary prolactotropes, thereby blocking the normal inhibitory effect of endogenous dopamine on PRL secretion. Dopamine, a key PRL inhibitory factor, is released from the hypothalamus and transmitted to the anterior pituitary gland by way of the hypophysial portal vessels. Opiates increase PRL levels by a central action that may involve dopamine neurons. Hyperprolactinemia linked to a microprolactinoma is clinically manifested by disturbed gonadal function, decreased libido, and various psychological disturbances (i.e., anxiety and depressive states) that can be ameliorated in part by the administration of dopamine agonists and/or serotonin antagonists. Such observations, as well as the phenomenon of pseudopregnancy, raise the possibility of PRL being involved in central nervous system (CNS) circuits beyond those involved in pregnancy and lactation.

Animal studies have indicated that PRL receptor-deficient mice have hyperprolactinemia, presumably due to the lack of feedback stimulation of TIDA neurons, and that the chronic elevation of PRL secretion occurs in rodents lacking a functional D_2 dopamine receptor. These findings suggest that PRL gains access to the CNS, presumably, via the arcuate–median eminence region, which has an incomplete blood–brain barrier. In experiments transplanting the pituitary to the kidney, there is excess PRL production because the pituitary grafts are remote from the site of dopamine release (in the hypothalamus).

The demonstration that there is PRL mRNA in the hypothalamus and other brain areas suggests that the behavioral effects assigned to PRL may relate to its presence in monoaminergic and peptidergic neurons. Immunohistochemical and *in situ* hybridization

studies revealed that PRL receptors are expressed in the arcuate, periventricular and medial preoptic nuclei both in male and nonpregnant female rats. In addition, during pregnancy and lactation PRL receptor immunoreactivity becomes demonstrable in the paraventricular nucleus (PVN) and ventromedial nucleus. The recruitment of a pleiad of neural structures during pregnancy and lactation assists in the provision of psychosomatic balance, needed for adequate nourishment of the mother and the fetus. The suppression of anxiety, reduced stress response, and expression of maternal behavior are linked to the activities of the PVN and medial preoptic nucleus, whereas the activation of the supraoptic ventromedial and arcuate nuclei are involved in regulating fuel and electrolyte balance.

During late pregnancy and lactation, the acute stress response is attenuated; it is hypothesized that this involves the PVN and the hypothalamic-pituitary-adrenal axis. The observation that PRL has dose-dependent anxiolytic effects in males and nonpregnant female rats (effects neutralized by the administration of antisense nucleotides against the PRL receptor), strongly suggests that PRL has a role as an endogenous anxiolytic. Stress-activated circuits enhance the release of cortical and medullary adrenal biological response modifiers – the glucocorticoids and catecholamines, respectively. When insulin-induced hypoglycemia (IIH) is used as a stressor, there may be an elevation of serum growth hormone (GH) and PRL. The blockade of PRL and enhancement of GH can be effected by dopamine agonists; an earlier clinical observation suggested that cyproheptadine blunts PRL release secondary to IIH, implicating indirectly the participation of serotonergic pathways. Whether the latter neurotransmitter is involved in the PRL release associated with the stress of anesthesia, sleep, nipple stimulation, gynecological exam, or exercise is unknown. Certainly, in contrast to dopamine and dopamine agonists, which block TRH-induced PRL release, serotonin (5-HT) antagonists are ineffective.

Beyond these observations about acute short-lived stress, there are some clinical observations that have stimulated research on psychological-psychiatric distress and PRL dynamics, independently of the activation of the hypothalamic-pituitary-adrenal axis. A woman on a birth-control pill in her teens developed breast engorgement and spontaneous bilateral galactorrhea, which reappeared 5 months later following general anesthesia for an orthopedic procedure. She had postpill amenorrhea with normoprolactinemia and no pituitary tumor; bromocriptine resulted in the cessation of galactorrhea and the reappearance of menses. This and other reports of dopamine agonist use in normoprolactinemic amenorrhea raise the question: what is normoprolactinemia in aberrant psychological situations? In earlier observations, during acrobatic flight, we observed a significant rise in serum PRL only in novices and not in trained pilots. Transient PRL elevations have been reported during medical interviews, before medical and/or competitive examinations, during argumentative conflicts in neurotic women, and during panic attacks (in the last, it was mentioned that the PRL increase was more consistent than that observed for cortisol, GH, or catecholamines).

In a study carried out in epileptics with a nonlateralized focus following galvanic stimulation of implanted electrodes (in the amygdala and hippocampus), we observed different GH and PRL responses in the corticomedial and basolateral amygdaloid regions. These and other reports (i.e., the different profiles of serum PRL among patients with unipolar vs. bipolar disorders) point to the involvement not only of the infundibular but also of mesocortical and mesolimbic dopamine systems controlling surges, circadian, and sleep-related PRL regulatory circuits. It was recently reported that hyperprolactinemic amenorrheic women had significantly higher symptom questionnaire scores on hostility, anxiety and depression than normoprolactinemic amenorrheic women or women with normal menstrual cycles. Thirty of these women met the *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn.; DSM-III) criteria for major depressive disorder, opening a PRL window for biological psychiatric research. An example along these lines is pseudopregnancy, which is considered a psychiatric condition in DSM-III (hysterical conversion neurosis); that is, the woman, firmly believing that she is pregnant, develops many of the symptoms and signs of pregnancy. In the absence of a prolactinoma, pseudopregnancy had been identified by Hippocrates: "If a woman is not pregnant and has not given birth, produces milk and her menstruation has stopped."

It is beyond any doubt that stress-induced hyper- or normoprolactinemic anovulation operates protectively for both the mother and the potential conceptus, for which the *milieu interieur* could be hostile. In the lactating, breast-feeding mother, the psychosomatic and immune well-being of the family (father, mother, and newborn) complex is preserved because breast-feeding inhibits ovulation and associated side effects. The immunoprotective actions of the somatolactotropic hormones and the maintenance of euthyroidism have been proposed as immunoprotective mechanisms ensuring reduced susceptibility to stress-induced disease. It is hoped that, in women with postpartum depression, hypnosis, psychotherapy, and the proper

choice of psychobiological response modifiers can be selected so that the hormonal-metabolic homeostasis is preserved.

Further Reading

Bridges, R. S., Numan, M., Ronsheim, P. M., et al. (1990). Central prolactin infusions stimulate maternal behavior in steroid-treated, nulliparous female rats. *Proceedings of the National Academy of Sciences USA* **87**, 8003–8007.

Dorshkind, K. and Horseman, N. (2000). The roles of prolactin, growth hormone, insulin-like growth factor-I, and thyroid hormone in immune system development and function: insights from genetic models of hormone and hormone receptor deficiency. *Endocrine Review* **21**, 292–312.

Dutt, A., Kaplitt, M. G., Kow, L. M., et al. (1994). Prolactin, central nervous system and behavior: a critical review. *Neuroendocrinology* **59**, 413–419.

Fava, G. A., Fava, M., Kellner, R., et al. (1981). Depression hostility and anxiety in hyperprolactinemic amenorrhea. *Psychotherapy and Psychosomatics* **36**(2), 122–128.

Grattan, D. R. (2002). Behavioural significance of prolactin signalling in the central nervous system during pregnancy and lactation. *Reproduction.* **123**(4), 497–506.

Hippocrates (1526) *Opera omnia*. Venice: Aldine Press.

Kelley, K. W. and Dantzer, R. (1991). Growth hormone and prolactin as natural antagonists of glucocorticoids in immunoregulation. In: Plotnikoff, N., Murgo, A., Faith, R. & Wybran, J. (eds.) *Stress and immunity*, pp. 433–452. Boca Raton, FL: CRC Press.

Pinter, E. J., Tolis, G., Guyda, H., et al. (1979). Hormonal and free fatty acid changes during strenuous flight in novices and trained personnel. *Psychoneuroendocrinology* J **4**(1), 79–82.

Sobrinho, L. G., Simoes, M., Barbosa, L., et al. (2003). Cortisol, prolactin, growth hormone and neurovegetative responses to emotions elicited during an hypnoidal state. *Psychoneuroendocrinology* **28**(1), 1–17.

Tolis, G., Dent, R. and Guyda, H. (1978). Opiates, prolactin, and the dopamine receptor. *Journal of Clinical Endocrinology and Metabolism* **47**(1), 200–203.

Tolis, G., Goldstein, M. and Friesen, H. G. (1973). Functional evaluation of prolactin secretion in patients with hypothalamic-pituitary disorders. *Journal of Clinical Investigation* **52**(4), 783–788.

Tolis, G. ., Labrie, F., Martin, J. B. and Naftolin, F. (1979). *Clinical neuroendocrinology: a pathophysiological approach*. New York: Raven Press.

Tolis, G., Stefanis, C., Mountokalakis, T. and Labrie, F. (1984). *Prolactin and prolactinomas: current developments*. New York: Raven Press.

Tolis, G. and Woods, J. F. (1980). Epilepsy and prolactin release in man: studies during spontaneous seizures or those induced via stereotactically placed electrodes in the human brain. (Abstract, p.465) Program of the 62nd Annual Meeting of the Endocrine Society. Washington.

Torner, L., Toschi, N., Pohlinger, A., et al. (2001). Anxiolytic and anti-stress effects of brain prolactin: improved efficacy ofantisense targeting of the prolactin receptor by molecular modeling. *Journal of Neuroscience* **21**, 3207–3214.

Pro-opiomelanocortin (POMC)

A B Bicknell
University of Reading, Reading, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A B Bicknell and P J Lowry, volume 3, pp 257–265, © 2000, Elsevier Inc.

Introduction
Discovery of POMC
Sites and Tissue-Specific Processing of POMC
Molecular Basis of Tissue-Specific Processing
Regulation of the Expression of the POMC Peptides
Posttranslational Modifications of POMC
Biological Activity of the POMC Peptides
Conclusions

Glossary

Adrenocorticotropin (ACTH)	A 39-amino-acid peptide released from the pituitary in response to stress that acts on the adrenal cortex promoting the synthesis and release of corticosteroids.
α-Melanocyte-stimulating hormone (α-MSH)	A peptide consisting of the first 13 residues of ACTH, produced mainly in the intermediate lobe of the pituitary. This peptide plays a role in the control of skin color in lower animals and in the control of feeding in mammals.
Immunoprecipitation	A purification process in which a specific antibody is used to capture a protein or peptide from a complex mixture (e.g., cell extract).
Prohormone convertases	A family of serine proteases that cleave proteins in the Golgi apparatus/secretory

granule before they are secreted into the circulation.

Pro-opiomelanocortin (POMC) A 30 kDa protein synthesized predominantly in the pituitary that is the common precursor to ACTH and several other biologically active peptides.

SDS-PAGE (Sodium dodecyl sulfate polyacrylamide electrophoresis) A method to separate proteins by size using electrophoresis through a gel matrix.

Introduction

Adrenocorticotropin (ACTH), like the vast majority of peptide hormones, is synthesized as part of a large, biologically inactive precursor protein that is then proteolytically processed to release the active peptides. ACTH is derived from pro-opiomelanocortin (POMC), a 30-kDa protein that is the common precursor to a number of biologically distinct peptides, an N-terminal region (N-POMC 1–76), ACTH, and β-lipotropin (β-LPH). These peptides can be further cleaved to give β-endorphin, corticotropin-like intermediate peptide (CLIP), and N-POMC 1–49. POMC also contains three copies of the melanocyte-stimulating hormone (α-, β-, and γ-MSH) sequence. The name of the protein thus encompasses the names of the peptides it encodes.

Discovery of POMC

The concept of hormone precursors was developed in the 1960s when it was shown that insulin was synthesized as a larger molecule, pro-insulin, and was released by limited proteolysis. During the 1970s evidence accumulated that implied a very close link between ACTH and β-LPH. The two peptides were released in approximately equimolar amounts from the pituitary, while immunocytochemistry showed the two peptides colocalized in the same cells. This evidence suggested that the peptides were derived from the same precursor, and in 1977 this hypothesis was proved to be correct when two groups showed, in two entirely different ways, that indeed ACTH and β-LPH were derived from the same molecule. Mains and Eipper used a pulse chase approach with the mouse pituitary tumor cell line AtT-20, which secretes large amounts of ACTH. They incubated the cells with radiolabeled amino acids (the pulse) and then washed the radioactivity away after a short period (the chase), thus labeling all proteins synthesized during the period of the pulse. By taking the cells at different times after the chase and immunoprecipitating the cellular protein using either an ACTH or β-endorphin antisera followed by SDS-PAGE analysis, they identified a common 31-kDa protein that became progressively smaller with time after the pulse until it became fully processed.

The second group, headed by Ed Herbert, also used the AtT-20 cell line together with a cell free translation system. Messenger RNA was purified from the cells and expressed in the cell free system. In this way a 28-kDa protein was identified that was recognized by both ACTH and β-endorphin antisera. The apparent difference in size was accounted for by the absence of glycosylation.

Although these experiments gave conclusive evidence that indeed ACTH and β-LPH were derived from the same protein, it was not until 1979, when the cDNA encoding POMC was cloned from both the bovine intermediate lobe and the mouse pituitary cell line AtT-20, that the true structural relationship between ACTH and β-LPH was revealed. Since then the POMC gene has been cloned and sequenced from a variety of organisms. In every species examined, it has been found that there is a single functional copy of the gene, and the overall gene structure, with the exception of the loss of γ-MSH in the salmon, has remained conserved throughout evolution. In the human it consists of three exons and two introns. Exons 2 and 3 are translated, with exon 2 coding for 44 amino acids; the remaining 223 residues are coded for by exon 3. The precursor molecule (pre-POMC) has a 26-amino-acid signal sequence followed by a 76-residue peptide (N-POMC or pro-γ-MSH) that contains one copy of the MSH sequence. This is followed by a joining peptide, ACTH, whose sequence encompasses the second MSH sequence (α-MSH), and finally β-LPH at the C-terminus, the sequence of which includes that for β-endorphin and the third MSH sequence. All the bioactive peptides within POMC are flanked by dibasic amino acid residues, e.g., Arg-Lys, Arg-Arg, Lys-Arg, and Lys-Lys. The structure of POMC and the major peptides released as a result of proteolytic cleavage are shown in **Figure 1**.

Sites and Tissue-Specific Processing of POMC

The main site of POMC expression is the anterior and intermediate lobes of the pituitary. In the anterior lobe, POMC represents approximately 1% of total protein, while it represents up to 10% in the intermediate lobe. Although these two tissues make up the principle sources of circulating POMC, it is also expressed and processed in several other tissues throughout the body, including the adrenal, gut, reproductive tract, placenta, leukocytes, spleen, lung, liver, thyroid, heart, skin, and brain.

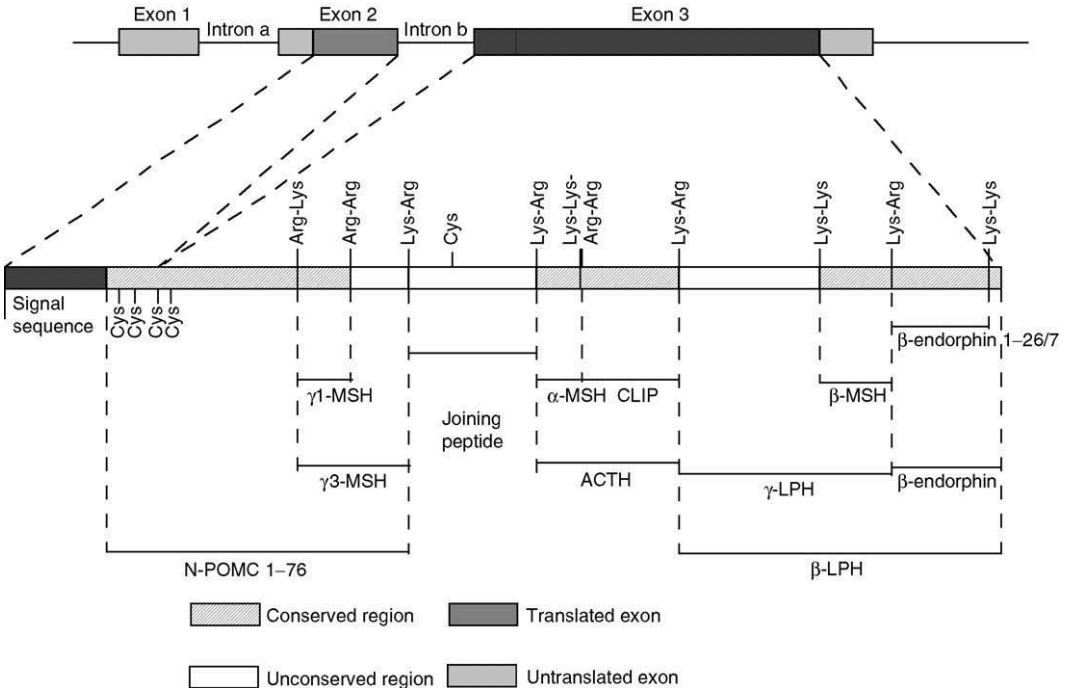


Figure 1 Structure of the POMC gene and the protein it encodes. Positions of the dibasic residues and the Cys residues are shown. Regions of the precursor conserved between species are also shown.

One advantage of producing peptides from a larger precursor is that the precursor can be cleaved in a tissue-specific manner to give a different set of products depending on tissue that the precursor is synthesized in. POMC is the classical example of this so-called tissue-specific processing, with the major sites of synthesis being the anterior and intermediate lobes of the pituitary. Classically, the pituitary can be divided into three distinct tissues:

1. The anterior pituitary or pars anterior contains a mixed population of cells, some of which, named corticotrophs, secrete POMC peptides that are involved mainly in the control of adrenal function.
2. The intermediate lobe or pars intermedia consists solely of melanotrophs that secrete POMC-derived peptides that are involved mainly with skin and hair pigmentation.
3. The posterior pituitary or pars nervosa stores and releases vasopressin and oxytocin produced in the hypothalamus.

In the rat these three tissues can be clearly defined, but in adult humans there is no distinct pars intermedia. It is present in fetal life but disappears soon after birth, leaving just a few scattered cells throughout the anterior lobe, giving the adult anterior lobe the collective name of pars distalis. The region between the anterior and posterior lobes is sometimes called the zona intermedia.

Processing of POMC in the Anterior Pituitary

Processing of the POMC in the anterior pituitary is summarized in Figure 2. The mouse anterior pituitary cell line AtT-20 has been used extensively as an *in vitro* model for processing in the anterior lobe. Using pulse chase experiments it was determined that the major products were a 16-kDa N-terminal fragment, ACTH, and β-LPH. β-endorphin can also be formed from sequential cleavage from β-LPH (not from direct cleavage of POMC) but is only a minor product.

Processing of POMC in the Pars Intermedia

Within the pars intermedia, processing is more extensive, with the sequential cleavage of peptides produced in the anterior lobe. ACTH is cleaved to α-MSH and CLIP, while β-LPH is virtually completely processed to β-endorphin, and γ-LPH β-endorphin (1–31) is cleaved at the C-terminus to give β-endorphin(1–27), β-endorphin(1–26), and a dipeptide glycylglutamine. The majority of the β-endorphin is N-acetylated. The N-terminal fragment N-POMC 1–76 is partially cleaved at an Arg-Lys site to give N-POMC 1–49 and Lys-γ3-MSH. This latter peptide contains an additional dibasic site (although not in all species) but does not appear to be cleaved to any significant extent.

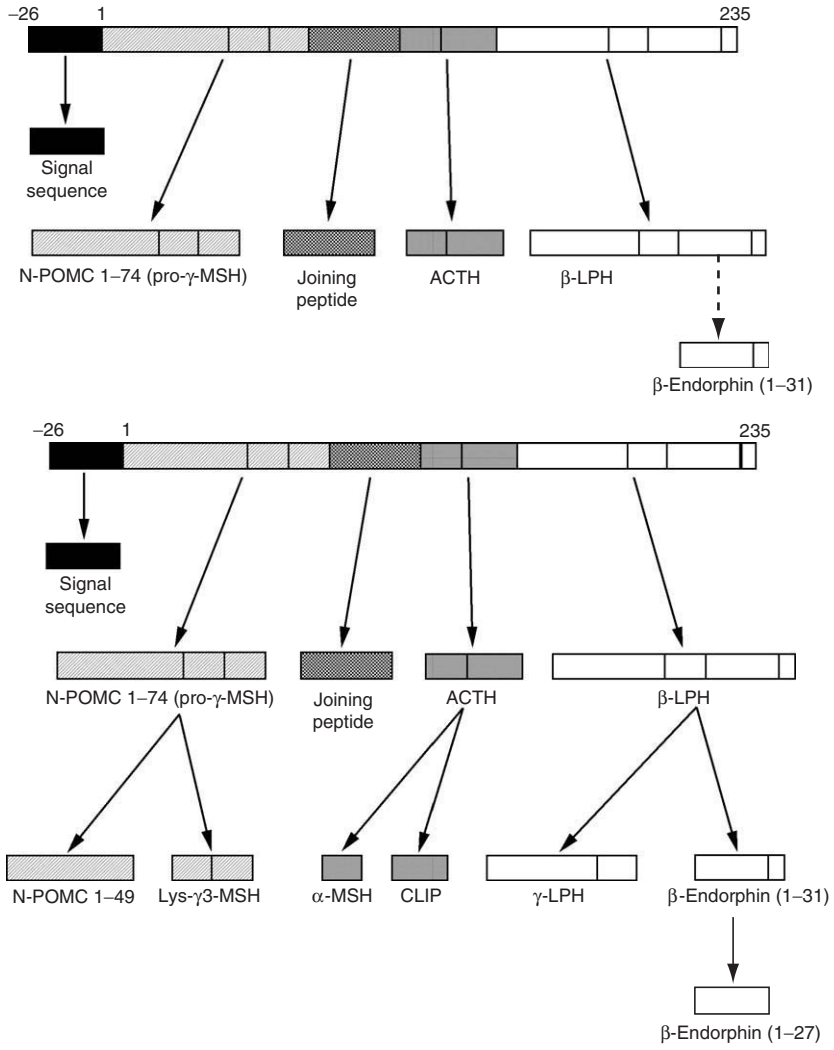


Figure 2 Processing of POMC in anterior and intermediate lobes of the pituitary. Dibasic cleavage sites are shown as vertical lines. Dashed arrow indicates incomplete processing.

Molecular Basis of Tissue-Specific Processing

POMC is itself biologically inactive and must be processed correctly to its biologically active peptides. During and after sorting and packaging of the POMC precursor into the regulated secretory vesicles, and prior to the release of the nascent secretory vesicles from the trans Golgi network, it is endoproteolytically cleaved. Cleavage is normally elicited at dibasic amino acid residues and follows a common route involving three different enzymes:

1. An endoprotease that cleaves on the C-terminal side of the basic amino acid pair.
2. A carboxypeptidase-B-like activity that removes the basic amino acid left after cleavage. Carboxypeptidase E (CPE) (also called CPH) specifically removes the basic residues from the C-terminus of a peptide.
3. A C-terminal amidating enzyme that catalyses the amidation of peptides terminating in glycine. The enzyme pepitidylglycine alpha-amidating monooxygenase (PAM) catalyses this reaction and utilizes a glycine situated on the N-terminal side of the dibasic cleavage site as the nitrogen donor for the amide group.

These processing events are carried out in an acidic environment within the trans Golgi network or within the immature secretory vesicles and are necessary for the optimum activity of the processing enzymes as well as to induce conformational changes in the protein to expose the basic amino acid residues for cleavage.

In most cases, specific processing of prohormones occurs on the carboxyl side of pairs of basic amino acids, usually Lys-Arg or Arg-Arg sites. Processing

can also occur at the tetrabasic sequence Arg-X-Lys/Arg-Arg and less frequently at monobasic sites. The enzymes that have implicated in the specific processing of POMC are known as prohormone convertases 1 and 2 (PC1 and PC2). These two enzymes are members of the subtilisin family of Ca^{2+}-dependent serine proteinases.

Identification of the Enzymes Involved in POMC Processing

During the early 1990s the identification of the prohormone converting enzymes was a major aim of researchers in the field. The breakthrough came with the cloning of a yeast gene called KEX2. The product of the KEX2 gene, known as kexin, is a protease that cleaves the yeast alpha factor mating pheromone. Cellular expression of the gene demonstrated that the enzyme belongs to the subtilisin family of Ca^{2+}-dependent serine proteinases. KEX2 has the ability to selectively cleave a number of mammalian prohormones, and this led to the hypothesis that similar mammalian counterparts existed.

The first mammalian gene to be identified as a potential processing enzyme came as a consequence of computer alignment of the amino acid sequences surrounding the serine and catalytically important asparagine of kexin against known protein sequences. By this method, a human gene of unknown function was identified, named *fur,* which lies upstream from the tyrosine kinase *fps/fes* proto-oncogene. The product of this gene, to become known as furin, was to become the first mammalian processing enzyme identified.

Using the sequence of the two eukaryotic proteases and several other subtilisin sequences, Smeekens and Steiner designed degenerate oligonucleotides corresponding to the conserved active site. By utilizing these in the polymerase chain reaction (PCR) with cDNA isolated from a human insulinoma library as template, they generated a fragment that was then used to probe that library. A full-length cDNA was identified that encoded for a protease that became known as PC2. Concurrently, Seidah and colleagues, using a similar procedure with the sequence of furin and a mouse pituitary library, identified two cDNA clones that were to become known as PC2 and PC1.

Elegant studies performed *in vitro* using cell-based assays and recombinant enzymes, together with studies looking at PC1 and PC2 knockout animals, have revealed the molecular basis of the tissue-specific processing observed in the anterior and intermediate lobes. In the anterior lobe, PC1 is the predominant enzyme present, and due to its more limited proteolytic activity, results in the generation of ACTH. However, in the intermediate lobe, both PC1 and PC2 are present, and their coordinate actions result in the generation of the smaller POMC peptides.

Cleavage of POMC occurs in a specific order and is summarized in Figure 3.

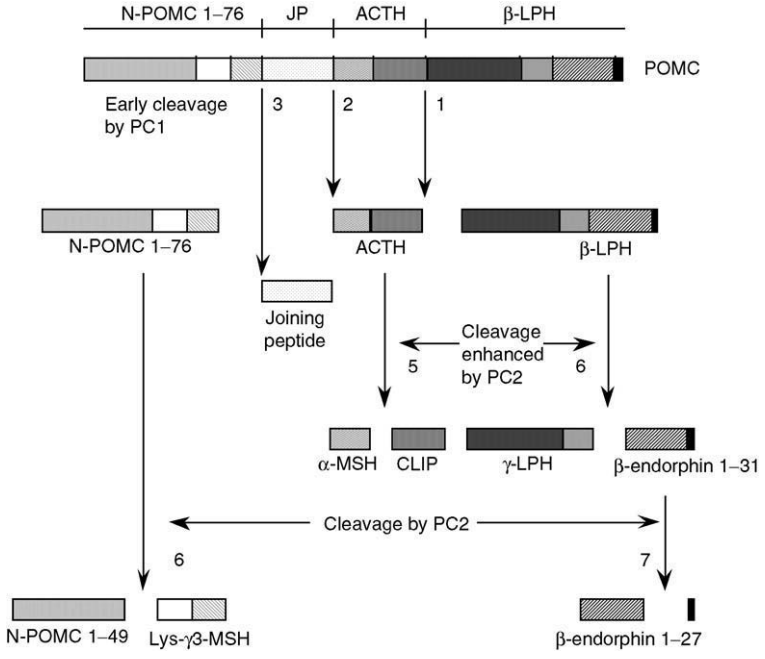


Figure 3 Processing of POMC by PC1 and PC2. The order of cleavage is dictated by the numbers showing that PC2 acts on the products of PC1 cleavage.

Regulation of the Expression of the POMC Peptides

As the processing of POMC occurs in a tissue-specific manner, it follows that the control of POMC expression and secretion of the derived products are also tissue specific.

Anterior Pituitary

With POMC produced in the pituitary, it followed logically that its secretion was under the control of factors released from the hypothalamus. In 1981 Vale identified the first of these as a 41-amino-acid peptide called corticotropin-releasing hormone (CRH). However, CRH does not stimulate ACTH release to the same extent as hypothalamic extracts, and it became clear that at least one other substance stimulated the release of ACTH. This second substance was identified by Lowry, McCann, Antoni and other investigators as arginine vasopressin, which potentiates the ACTH-releasing activity of CRH. Stimulation by CRH results in an increase in POMC mRNA levels as well as an increase in secretion of peptides. The two main factors are also potentiated by the action of several other weak secretagogues, namely, epinephrine, norepinephrine, oxytocin, and angiotensin II.

Observations after treatment with exogenous glucocorticoids or adrenalectomy have long implicated a role for glucocorticoids in the control of POMC peptide secretion. They exert their effects both at the hypothalamus, by decreasing CRH secretion, and at the corticotroph, by decreasing POMC synthesis.

Pars Intermedia

As would be expected, the regulation of POMC expression and secretion in the pars intermedia is quite distinct from that in the pars anterior. Glucocorticoids have no influence on POMC expression; although receptors are present, they do not seem to elicit any control over POMC gene expression.

POMC gene expression is enhanced by stress CRH and β-adrenergic agonists All exert their actions via cAMP.

Dopamine decreases POMC expression by binding to D2 receptors, which are thought to block the activation of adenylate cyclase, although there may be more direct effects.

Posttranslational Modifications of POMC

POMC and its peptides undergo a number of different modifications following POMC's synthesis and removal of the signal peptide.

Disulfide Bridge Formation

Human POMC contains five Cys residues, one more than every other species. Four occur close to the N-terminal at positions 2, 8, 20, and 24, while the fifth occurs in the joining peptide. The four N-terminal Cys residues form intramolecular disulfide bridges between 4 and 24 and 8 and 20. The fifth Cys, contained within the joining peptide and unique to humans, leads to its secretion as a homodimer, which may indicate that human POMC is produced as a dimer.

Glycosylation

In all species Asn 65 is N-linked glycosylated in the consensus sequence Asn-X-Ser. O-linked glycosylation also occurs at Thr 45 in the human and porcine pituitary.

Phosphorylation

Serine 31 of ACTH is phosphorylated in the rat, human, and mouse, although its presence has no effect on biological activity.

Acetylation

α-MSH and β-endorphin are found as des-, mono- and diacetylated forms, although the precise form varies between tissue and species. The N-terminal Ser of α-MSH may be acetylated through either the hydroxyl or amide group, acetylation increasing both its biological activity and half life, in contrast to β-endorphin, which reduces its biological opiate activity.

Biological Activity of the POMC Peptides

The peptides contained with the POMC precursor have a wide spectrum of biological activity that is briefly described in the following sections.

Adrenocorticotropin (ACTH)

The main role of ACTH is to promote steroidogenesis within the adrenal cortex. It exerts its actions by binding to the melanocortin 2 receptor (MC2R), leading to a rise in the levels of cAMP and activation of protein kinase A, resulting in secretion of glucocorticoids, androgenic steroids, and, to a lesser extent, mineralocorticoids. The acute effect of ACTH is to increase the rate of conversion of cholesterol to pregnenolone, the first and rate-limiting step in the steroidogenic pathway. The chronic effects of ACTH result in an upregulation of the expression of most of the enzymes involved in the steroidogenic pathway.

ACTH has also been implicated in the maintenance of adrenal weight. The evidence for this is based heavily on the observation that long-term glucocorticoid therapy results in adrenal atrophy. However, since all POMC peptides are downregulated during such treatment and administration of ACTH alone

does not totally reverse the atrophy, it is clear that other POMC peptides play a role in the control of adrenal weight.

N-Terminal Peptides and Adrenal Growth

The role of the N-terminus of POMC in adrenal growth was originally proposed in the early 1980s, when a series of elegant studies demonstrated that N-terminal peptides could stimulate mitogenesis in adrenal cell cultures. It was also shown that immunoneutralization of N-POMC peptides resulted in the inhibition of the compensatory growth response observed following unilateral adrenalectomy.

However, N-POMC 1–76 is itself inactive, but peptides derived from its N-terminus, not containing γ-MSH, are potent adrenal mitogens. Since these shorter peptides are not generated during normal pituitary processing, it was proposed that N-POMC 1–76 is cleaved after secretion. Considerable weight has recently been added to this argument with the identification of an adrenal protease that appears to cleave N-POMC to the shorter forms.

Melanocyte-Stimulating Hormone

There are three copies of the MSH sequence in the POMC precursor: α-MSH in ACTH, β-MSH in β-LPH, and γ-MSH in N-POMC 1–76. The biological role of the MSH peptides in lower vertebrates such as frogs was known before the discovery of POMC. α-MSH produced by the pars intermedia causes the darkening of skin to match a darker background. Both α-MSH and β-MSH share the common core sequence Met-Glu-His-Phe-Arg-Trp-Gly, which is essential for biological activity. γ-MSH contains the sequence Met-Gly-His-Phe-Arg-Trp-Asp and was named due to its similarity to the other MSH sequences. However, it has no melanotropic activity.

The role of the MSH peptides in skin pigmentation in humans is not clear. Injection of the peptides into humans does result in skin darkening. Also, in Nelson's syndrome, a condition in which there are high levels of POMC peptides, skin darkening is observed. However, in the normal human this role is debatable. Humans do not have any pars intermedia tissue, and certainly there are no detectable amounts of α-MSH in the periphery, although mRNA encoding POMC can be detected in human skin.

Role of MSH Peptides in the Regulation of Feeding

Mammals regulate their energy stores relatively constantly despite a large variation in both food availability and physical activity. The control of energy expenditure has been shown to be the result of a complicated endocrine loop involving a number of neuropeptides, including the melanocortins.

POMC is synthesized in the hypothalamic arcuate nucleus, where it is processed in a manner similar to that observed in the intermediate lobe. It has been generally assumed that α-MSH is the most important peptide since it is the ligand with highest affinity for the two melanocortin receptors – the MC3R and MC4R expressed in the hypothalamus. Genetic and pharmacological disruption of this loop causes obesity in both humans and rodents. The POMC knockout mouse is hyperphagic and obese, whereas infusion of α-MSH (or synthetic agonists) results in anorexia and weight loss.

The complete subtle architecture of this system is still far from clear, but the system's potential as a therapeutic target for the treatment of obesity has made it the focus of intense research over the past few years.

γ-MSH and Steroidogenesis

γ-MSH is generated from N-POMC in the intermediate lobe of the pituitary. It contains a slightly modified core sequence, Met-Gly-His-Phe-Arg-Trp-Asp, and has a biological role distinct from that of α-MSH.

Studies performed by Alastair Brownie in the early 1980s identified γ-MSH as a regulator of hormone sensitive lipase (HSL). This enzyme is found in steroidogenic tissues and breaks down cholesterol esters to free cholesterol, which then feeds into the steroidogenic pathway. Several groups demonstrated that although not directly steroidogenic, this peptide can synergize with ACTH and result in an increase in steroid output. The receptor through which this peptide acts is unclear, since although the MC3R has been shown to be the only melanocortin receptor with significant affinity for γ-MSH, its expression profile and pharmacology suggest that it does not play a role in this system.

Joining Peptide

The joining peptide has no known biological function. Its role, like the C-peptide of insulin, is probably only to provide a spacer between the different peptides. One interesting feature of human joining peptide is that it uniquely contains a cysteine that has been implicated in dimer formation.

β-Endorphin

β-endorphin is one of several endogenous opioids that exert their actions by binding to receptors within the central and peripheral nervous systems. It is part of a larger family whose members include dynorphin and

the enkephalins. The opioids bind to three receptors known as the μ, κ, and δ opioid receptors. The receptors derive their names from the opiate agonists that bind to them: morphine and ketocyclazocine for the μ and κ receptors, and the vas deferens for the δ receptor, where it is located. The receptors differ in their affinity for the different opioids, although not one is specific for a single ligand. β-endorphin and Met-enkephalin bind to the μ and δ receptors, while dynorphin binds to the κ receptor. All the receptors are G-coupled and recognize a common motif consisting of a protonated amine adjacent to an aromatic amine, for example, the tyrosine residue at the N-terminus of the enkephalins. The opioid actions of β-endorphin are dependent on its C-terminal residues. Removal leads to a decrease in both its analgesic properties and binding affinity for the receptors. Thus, the shorter endorphins lack opioid activity, as does N-acetylation of the peptide. The N-terminal portion of β-endorphin is required for binding to the opiate receptors, whereas the C-terminal residues are thought to enhance the binding affinity.

Conclusions

Most peptide hormones are produced as inactive precursor molecules that are subsequently cleaved at dibasic residues to release the active peptides. As the first such multihormone to be discovered, over a quarter of a century ago, POMC has been extensively studied and characterized. Its role in the control of adrenal function by the actions of ACTH on steroidogenesis, as well as N-POMC on the control of adrenal tonic state, are central to the stress axis as a whole.

See Also the Following Articles

Adrenocorticotropic Hormone (ACTH); Obesity, Stress and; Peptides.

Further Reading

Bertagna, X. (1994). Proopiomelanocortin-derived peptides. *Endocrinology and Metabolism Clinics of North America* **23**, 467–485.

MacNeil, D. J., Howard, A. D., Guan, X., et al. (2002). The role of melanocortins in body weight regulation: opportunities for the treatment of obesity. *European Journal of Pharmacology* **450**, 93–109.

Raffin-Sanson, M. L., de Keyzer, Y. and Bertagna, X. (2003). Pro-opiomelanocortin, a polypeptide precursor with multiple functions: from physiology to pathological conditions. *European Journal of Endocrinology* **149**, 79–90.

Smith, A. I. and Funder, J. W. (1988). Proopiomelanocortin processing in the pituitary, central nervous system and peripheral tissues. *Endocrine Reviews* **9**, 159–179.

Steiner, D. F. (1998). The proprotein convertases. *Current Opinion in Chemical Biology* **2**, 31–39.

Prostaglandins

S Moshonov and U Zor
Weizmann Institute of Science, Rehovot, Israel
Z Naor
Tel Aviv University, Tel Aviv, Israel

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Moshonov, U Zor and Z Naor, volume 3, pp 266–272, © 2000, Elsevier Inc.

Introduction
Characteristics of Prostaglandins
Prostaglandins in the Hypothalamic-Pituitary-Adrenal Axis
Stress and Antistress: Prostaglandins versus Glucocorticosteroids

Glossary

Adrenals	The organs that cap the kidneys. The outer layer, the cortex, synthesizes the corticosteroids, whereas the inner layer, the medulla, synthesizes epinephrine.
Cytokines	Regulatory proteins, such as the interleukins and lymphokines, that are released by cells of the immune system and act as intercellular mediators in the generation of an immune response.
Glucocorticosteroids	Anti-inflammatory steroids; also called glucocorticoids. Cortisol and corticosterone are natural corticosteroids secreted by the adrenal cortex and have strong glucocorticoid activity.
Hypothalamus	A specialized area in the brain that regulates automatic body functions such as temperature, sleep, and water balance. It coordinates the endocrine system.

Noradrenergic cells	Nerve cells that release norepinephrine.
Nonsteroidal anti-inflammatory drugs (NSAIDs)	Drugs (e.g. aspirin, indomethacin) that act by inhibition of the cyclo-oxygenases (COX-1 and COX-2) and hence prevent prostaglandin synthesis.
Phospholipids	Lipid (fat) molecules with a nonlipid phosphate head. Phospholipid bilayers are the basic components of cell membranes.
Phosphorylation	The process of adding phosphorus to a dormant enzyme, thereby activating it. Occasionally, phosphorylation deactivates an active enzyme.
Pituitary	The master endocrine gland that lies just beneath and is attached by the pituitary stalk to hypothalamus. The pituitary controls major endocrine systems, growth, body water and electrolytes, lactation and pasturition. The pituitary gland is under central nervous control, which in the case of the anterior pituitary gland is mediated by hypothalamic neurohormones transported from the hypothalamus by the hypohysial portal vessel system.
Prostaglandins (PGs)	A family of hormones whose chemical structure is identical, except for small changes in the pentane ring; these differences provide the distinguishing nomenclature (prostaglandin D, E, F, and I and thromboxane A). The subscript in the name (e.g., PGD_2) refers to the number of unsaturated carbon bonds.
Stressors	Any agent that induces stress in the intact animal or in cell or tissues in culture.

Introduction

Stress, generally classified in terms of emotional or physical, presents a very clear pattern of pathophysiological changes. These are the consequences of stimulation by the many biologically active agents – hormones, neurotransmitters, and cytokines – generated by the nervous, endocrine, and immune systems aroused by stress. To a large extent, the systems are interdependent, and their complex networking is still far from being completely understood. The body maintains its healthy state by continuously readjusting to the changing environment, a process known as homeostasis. Stressors push the homeostatic mechanisms beyond normal limits, and the body enters a new, temporary state called stress.

The participation of the main agents of stress, adrenocorticotropic hormone (ACTH), corticosteroids, and catecholamines (CAs), has been long established. PGs, which are local rather than transported hormones, are not among these main stress hormones that are poured out immediately in response to stress. However, cells from stressed animals produce more PGs than controls do, and PG levels in serum, spinal fluid, and the hypothalamus are elevated during stress. The role of PGs in stress is still largely undefined, but they appear to be involved in at least two stages (Table 1): (1) the initiation of the stress response in the brain and (2) the expression of some symptoms of stress.

Table 1 Stages in the response to a stress signal that may be affected by prostaglandins

Stages of stress response	*Prostaglandin link*
Recognition of stress signal	Unclear
Initiation of stress response	Positive
Transmission and modulation of stress response	Possible
Expression of stress symptoms	Positive

Characteristics of Prostaglandins

Background

As first reported in the 1930s, PGs were the most powerful, ubiquitous, diversified, and mysterious biological substances yet discovered. Some 30 years later, they were finally unmasked by the Swedish medicinal chemists Süne Bergström and Bengt Samuelsson, who in 1982 were awarded the Nobel Prize in Physiology with the British pharmacologist John Vane. They isolated, identified, and characterized the PG family, establishing unsaturated lipids derived from the cell membrane as an unexpected class of highly sensitive bioactive regulators. Thanks to Vane and colleagues, the mechanism of action of aspirin, the first of the manufactured NSAIDs and one of the most successful drugs ever to be marketed, was now, some 70 years later, finally elucidated, proving to be a specific inhibitor of PG synthesis.

Synthesis of Prostaglandins

Cell membranes, comprising largely phospholipids, are themselves the major source material for PGs and other biologically reactive eicosanoids (20-carbon unsaturated carboxylic acids) such as leukotrienes (LTs).

Phospholipase A_2 Arachidonic acid (AA), an unsaturated fatty acid, is detached from membrane phospholipids by the enzyme phospholipase A_2 (PLA_2), transformed by cyclo-oxygenases (COXs) into the unstable endoperoxide PGH_2, and then metabolized by the various synthases to PGE_2, $PGF_{2\alpha}$, PGD_2, prostacyclin (PGI_2), or TXA_2 (Figure 1). The superfamily of the PLA_2 enzymes consists of 14 groups (I–XIV)

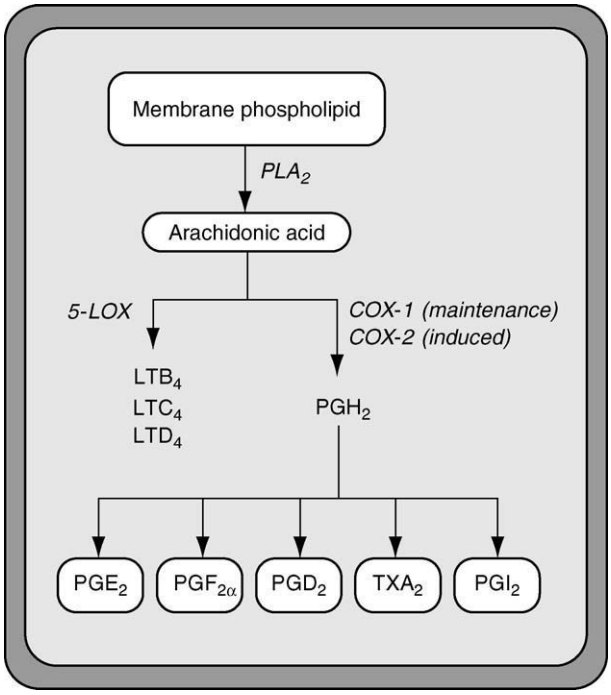


Figure 1 Prostaglandins and their synthesis. COX, cyclo-oxygenase; LOX, lipoxygenase; LT, leukotriene; PG, prostaglandin; PGI_2, prostacyclin; PLA_2, phospholipase A_2; TXA_2, thromboxane A_2.

and at least 19 members. The low-molecular-mass (~12- to 19-kDa) secretory PLA_2 ($sPLA_2$) proteins are all related evolutionarily, share a common catalytic histidine, and fall into the groups I–III, V, and IX–XIV. A high-molecular-mass (60- to 110-kDa) family of calcium-dependent (requiring micromoles of Ca^{2+} for activation) cytosolic PLA_2 ($cPLA_2$) proteins use a catalytic serine and have a high specificity to AA. Some are activated by Ser505 phosphorylation, use a novel catalytic dyad of Ser228 and Asp549, and are designated group IV (A, B, C or α, β, γ $cPLA_2$). The calcium-independent PLA_2 ($iPLA_2$) (group VI) contains eight ankyrin repeats, has a lipase consensus sequence with an active serine that serves as a nucleophile, plays a role in phospholipids' remodeling and homeostasis via the production of lysophospholipids, and falls into the subgroups (A-1, A-2, and B). Groups VII and VIII PLA_2 are platelet-activating factor (PAF)-acetylhydrolases that cleave PAF to form lyso-PAF and a free acetyl group. mRNA expression of $cPLA_2$ is enhanced by pro-inflammatory cytokines and inhibited by glucocorticoids (GCs).

An alternative and equally important route for AA is metabolism by lipoxygenases (LOXs) to the exceedingly potent LT derivatives (Figure 1). Like PGE_2 and $PGF_{2\alpha}$, LTs mediate inflammatory and allergic reactions, but their contribution to stress is largely uncharted. Other branches of the PG–LT family are also not discussed in this article; they appear to be less bioreactive and less is known in general about their contribution to pathophysiological mechanisms.

Cyclo-oxygenases The COX enzyme was first known as PGH_2 synthase (PGHS) because of its ability to catalyze the first two reactions in PG synthesis from arachidonate to PGH_2. COX was first isolated in 1976, and its two isoforms, COX-1 and COX-2, were discovered around 1992. The two isoforms (71 kDa), are encoded by different genes, are structurally similar, are composed of some 600 amino acids, and catalyze the same substrate. The enzymes contain a heme group that binds molecular oxygen. COX-1 is mainly found in the endoplasmic membrane, whereas COX-2 is found also in the nuclear envelope. The active sites of the two enzymes differ, which may explain why the enzymes are inhibited by different drugs. COX-1 is constitutive (i.e., always present in almost all cell types) and is associated with normal physiological PG activities, such as protection of the kidney and the stomach lining from damage and maintenance of vascular homeostasis. However, COX-2 is induced by stimuli, both stressful ones such as inflammatory agents and natural ones such as growth factors; it is also found constitutively in a few organs, one of which is the brain. It is not, however, expressed in nerve cells that synthesize the PGs responsible for stress-induced fever, hyperalgesia, and sleepiness. The origin of this stress-induced COX-2 is the endothelial cells lining the blood vessels in the brain, which are accessed easily by the PG-stimulating blood-borne cytokines. Anti-inflammatory GCs, complexed with their specific receptors, penetrate the nucleus, bind to DNA at the hormone response element sequence, regulate transcription, and preferentially suppress the induction of COX-2.

Nature and Functions of Prostaglandins

Prostaglandins are very potent local hormones that are synthesized rapidly by a wide variety of cells in the body. PGs are released into the extracellular space by specific PG transporters. Once released, PGs bind to their cognate receptors, which belong to the G-protein-coupled receptor (GPCR) superfamily. There are four major receptor subtypes for PGE_2 (EP1, EP2, EP3, and EP4), which are encoded by separate genes. There is a major receptor for $PGF_{2\alpha}$ (FP), for PGD_2 (DP), and for PGI_2 (IP). The various PG receptors differ in their signaling pathways. Sequence homology is found between receptors that share a signaling cascade rather than a ligand. Therefore, the group IP, DP, EP2, and EP4, which act via Gs to elevate cAMP levels, are known as the relaxant receptors because of

their common effect on smooth muscle relaxation. The contractile receptors TP, FP, and EP1 mobilize calcium via Gq and induce smooth muscle contraction. PGs, therefore, act in an autocrine/paracrine manner via cognate GPCRs. Within a short time, usually less than 30 min and in the case of thromboxane A_2 (TXA_2) a mere 30 s, PGs are inactivated (90%) to oxidized metabolites when passing through the lungs; very little PG is normally found in the circulation.

Pathological functions of prostaglandins and connections with stress PGs almost always act in the vicinity of their synthesis. The blood-borne, pro-inflammatory cytokines interleukin (IL)-1, IL-6, and tumor necrosis factor (TNF)-α are released from leukocytes in the blood and macrophages in tissues and are among the agents that locally trigger PG synthesis.

- Inflammation. Traumatic, inflammatory, immune, or other foreign stimuli trigger the activation and recruitment of PGs and LTs. They are part of the inflammatory reaction, promoting leukocyte infiltration, edema, and local vasodilation. Stress generates the same arachidonic AA metabolites and cytokines as inflammation.
- Fever. The elevation of body temperature, but not the maintenance of core temperature, is controlled by PGE_2 in the preoptic area of the hypothalamus. Stress-induced fever is reduced by the prior administration of aspirin-like drugs, such as indomethacin. The stimulation of PGE_2 synthesis in the endothelial cells of hypothalamic blood vessels by stressors such as lipopolysaccharide (LPS), which mimic the symptoms and the stress of bacterial infection in experimental animals, elevates body temperature. Fever also accompanies the stress of surgery (trauma), pain, and cage switching in rats, restraint in pigs, and open field (but not the stress of restraint) in rats or pigs being transported or frustrated by a lack of reward after exhibiting reward-training behavior.
- Hyperalgesia. Heightened sensitivity to pain, or hyperalgesia, is another stress symptom related to PGs. Pain is processed along the spinothalamic tract, reaching the thalamus, limbic system, and frontal cortex before descending through moderating pain-inhibitory nerve pathways. Stress-induced COX-2 is most evident in the spinal cord. Animal models of inflammatory stress and pain, such as carrageenin injected into rat foot pads, show that PGs are generated locally at the sensory nerve endings and in the spinal cord and cause hyperalgesia. COX-2 inhibitors (such as NS398 and DuP 697) and mixed COX-1 and COX-2 inhibitors (such as aspirin and indomethacin), applied directly to the spinal fluid, are analgesic.
- Dysmenorrhea. Dysmenorrhea is a condition with painful menstrual periods with uterine hypercontractility. Increased levels of PGE_2 and $PGF_{2\alpha}$ were found in the menstrual fluid of women with dysmenorrhea, and COX enzyme inhibitors alleviate the pain.
- Endometriosis. Endometriosis occurs when endometrial tissue is found in extrauterine sites; it is associated with pelvic pain and infertility. The increased production of PGE_2 and $PGF_{2\alpha}$ and overexpression of COX-2 are thought to be associated with endometriosis. Estrogen and angiogenic factors, such as vascular endothelial growth factor (VEGF), may be involved in the proliferative responses induced by the PGs in endometriosis.
- Cancer. Various studies have indicated a role for the COX enzymes, PGs, and PG receptors in the pathology of cancers. The overexpression of the COX enzymes and PG receptors and the increased production of PG were found in various cancers. The studies suggested the use of COX enzyme inhibitors in treating some cancers. Because COX-1 is regarded as a constitutive enzyme, whereas COX-2 is the regulated enzyme, attention has been paid to the development of selective COX-2 inhibitors.

Some physiological functions of prostaglandins with connections to stress The various PGs have diverse and frequently opposing actions (Table 2). PGs used in nonpathological functions are synthesized mainly by the constitutive (maintenance) enzyme COX-1.

- Blood pressure. PGE_2 and prostacyclin (PGI_2) are vasodilators and have a role in normal blood pressure control. However, TXA_2 is a strong vasoconstrictor. Blood pressure rises in stress and is not affected by PG inhibitors.
- Gastric ulcers. In the gastrointestinal system, PGE_2 and PGI_2 protect the stomach lining by inhibiting gastric acid secretion and promoting the healing of ulcers. Prolonged use of aspirin or other NSAIDs to inhibit PG synthesis. Gastric ulcers are enhanced.
- Preterm birth. The presence of PGE_2 and $PGF_{2\alpha}$, both of which contract the pregnant uterus and are produced by uterine and fetal tissue, is essential for the survival of the fetus. During maternal or fetal stress, these PGs are induced and can lead to preterm birth.
- Sleep. A central function of PGD_2 is thought to be sleep induction. Its metabolic enzyme PGD_2 synthase is located in the cerebrospinal fluid and the meninges and choroid plexus, nonneuronal cells lining the whole central nervous system (CNS).

Table 2 Diversity of physiological and pathological effects of the prostaglandins[a]

	PGD_2	PGE_2	$PGF_{2\alpha}$	PGI_2	TXA_2
Blood pressure	Down	Down	No change	Down	Up
Airway	Constrict	Dilate	Constrict	Dilate	Constrict
Uterus (nonpregnant)		Relax	Contract		
Uterus (pregnant)		Contract	Contract		
Kidney (GFR)	Increase	No change		No change	Reduce
Stomach secretions		Inhibit	Stimulate	Inhibit	
Gut motility		Increase	Increase	Increase	Increase
Platelet aggregation			Inhibit		Stimulate
Sleep	Induce				
Fever		Induce			
Pain		Lowers threshold		Lowers threshold	

[a]Chart is not comprehensive.

Sleepiness is a stress response, and PGD_2 could be involved in this process.

Stimulation of Prostaglandin Synthesis in Stress

Most stressors stimulate PG synthesis. Increased plasma (or medium) PGE_2 was induced by heat, pain, trauma (bone fracture), exercise, LPS, restraint, water immersion, and cage switching. A fall in PGE_2, however, was associated with cold stress in rats and captivity in frogs. In these captive frogs, $PGF_{2\alpha}$ increased, as it also did in individuals subjected to the stress of surgery, students taking examinations, and cells that were heat-stressed. PGI_2 was also increased by examination and surgery stresses, and it was removed from the aorta of immobilized rats.

In the brain of stressed animals, IL-1-stimulated PGE_2 releases the corticotropin releasing hormone (CRH) in the hypothalamus, thus activating the hypothalamic-pituitary-adrenal (HPA) axis. Most studies report that PGE_2 is synthesized by the cytokines in the brain; $PGF_{2\alpha}$ and PGD_2 are also mentioned in such studies. PGs induced by IL-1, IL-6, and TNF-α in the preoptic area also bring about fever and cause hyperalgesia via the spinal cord (and at sensory nerve endings), whereas PGD_2 in the cerebrospinal fluid and meninges induces sleep. The sources of these PGs are not necessarily neuronal but, rather, originate from the blood vessels, microglia, or other cells lining the brain.

Assay for Prostaglandins in Stress

Types of stress and measurement of prostaglandins For pharmacological purposes, there is no gross difference between emotional and physical stress. Cell systems galvanized by the body to express and to cope with the stress of, for example, bereavement (emotional or psychological) are ultimately the same ones used in the stress caused by an illness (internal physical) or a supreme athletic effort (external physical). An element of psychological stress generally accompanies a physical stress.

Experimental models of stress have been devised to mimic human stress-provoking situations, in which animals, or cells or tissues in culture, are subjected to one of a number of unpleasant procedures. In these models, stress behavior, in the presence and in the absence of drugs that modify PG synthesis, is observed and the biochemical changes are monitored. Alternatively, tissue samples are assayed for PGs, PLA_2, or COX-2 before and after provoking stress. PGs are commonly assayed using radioimmunoassay techniques, and enzymes are assayed by measuring their activity on labeled substrates or measuring their mRNA levels. Increases in enzyme activity (PLA_2 and COX) and a rise in concentration of a specific PG have been consistently reported in many different stress states.

Inhibitors of prostaglandins Prostaglandin synthesis is inhibited by two groups of anti-inflammatory agents: NSAIDs and GCs (the very same hormones elicited by stress). NSAIDs interfere only with COX activity; the prime example is aspirin, although indomethacin is routinely used in experimental systems. These drugs inhibit COX-1 and COX-2 to varying degrees. COX-2-selective inhibitors (COXibs) such as Celecoxib (Celebrex) and Rofecoxib (Vioxx), developed in the 1990s, are relatively large inhibitors that penetrate the active site of COX-2 but not that of COX-1. They offered much promise when they first appeared on the market, but were found to be associated with a potential risk of myocardial infarction and sudden cardiac death. More recently, attention has been given to the pharmacology of the PG receptors. Efforts are being made to develop receptor-specific antagonists that will inhibit the biological action of the PG rather than their synthesis. GCs, such as dexamethasone, depress the whole AA metabolite network by partial inhibition of PLA_2

synthesis and control stress-induced PGs, in particular by preventing cells from inducing COX-2. The very strong anti-inflammatory action of GCs is due to their wide-ranging anti-inflammatory base. The GC receptor complex interacts directly with and regulates the gene transcription of many inflammatory enzymes and proteins, including COX-2. Inhibitors of PG have been widely employed in the stress models to determine the contribution of PGs to the stress symptoms. This use of PG inhibitors as indirect proof has its limitations because the NSAIDs that can distinguish between the various PGs have not yet been developed and GCs suppress a whole spectrum of agents provoked by stress.

Prostaglandins in the Hypothalamic-Pituitary-Adrenal Axis

Stress Hormone Flow

The perception of distress is instantly transmitted via one or more neural circuits, each with its own neurotransmitter system, to the median eminence (ME) in the hypothalamus. The production of CRH, the major stress hormone released by the hypothalamus, and arginine vasopressin (AVP) is intensified immediately and stimulates the anterior pituitary to release ACTH into the circulation. The adrenal cortex is stimulated by ACTH to produce cortisol and corticosterone, powerful GCs that, by suppressing the production of pro-stress agents, including PGs, keep the sharp stress response from running fatally out of control (Figure 2).

The sympathetic nervous system, also activated in stress, releases CAs into the circulation – norepinephrine (NE) and epinephrine (EPI) – setting the body in the stress-alert mode for flight or fight. CAs stimulate PG synthesis in many tissues, including the brain.

Prostaglandins in the Initiation of Stress Response

Prostaglandin activity has been located in the HPA axis, prior to CRH release. The stress-related PG presence has not been confirmed at other points along the axis. One of the CRH-stimulating neural circuits is noradrenergic. These neurons originate in and/or project to the paraventricular nucleus (PVN) of the hypothalamus and are stimulated by IL-1- and IL-6-induced PGE_2. The rat hypothalamus can synthesize PGE_2, $PGF_{2\alpha}$, PGI_2, and TXA_2, but, at least *in vitro*, only PGE_2 is produced in response to IL-1 and IL-6, although PGD_2, $PGF_{2\alpha}$, and TXA_2 also induce CRH release. The noradrenergic circuit that stimulates the HPA axis can stimulate PG synthesis and is inhibited by indomethacin. PGs also seem to be able to directly induce CRH or AVP release. Thus, PGs may activate the HPA axis through the direct stimulation of CRH or AVP neurons or via noradrenergic neurons in the PVN (Figure 3).

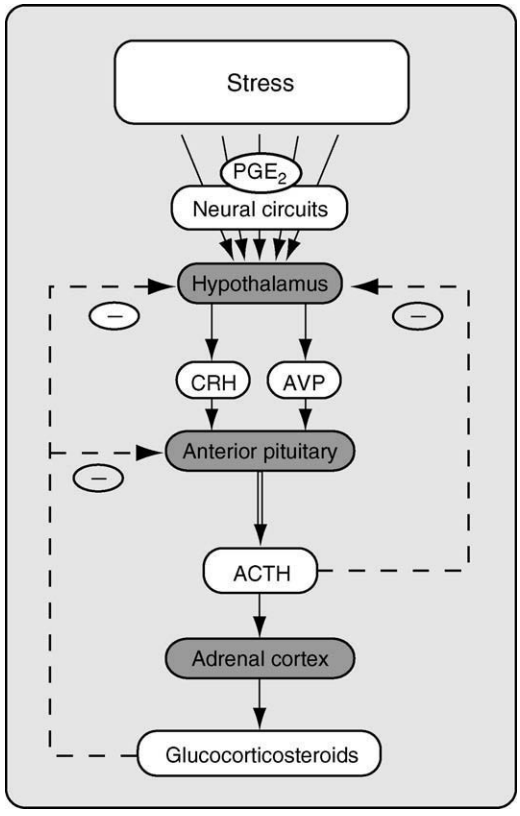


Figure 2 Activation of the hypothalamic-pituitary-adrenal (HPA) axis by stress. Release of adrenocorticotropic hormone (ACTH) is under the negative control (–) of the corticosteroids it induces. Prostaglandin E_2 (PGE_2) stimulates the release of corticotropin releasing hormone (CRH) or arginine vasopressin (AVP) from the hypothalamus and mediates the noradrenergic activation of the HPA axis.

Stress and Antistress: Prostaglandins versus Glucocorticosteroids

The stress-induced production of pathological PGs, together with anti-stress GCs, raises the question of their mutual operation during stress. The GC surge elicited by stress effects an inbuilt antistress facet to limit some aspects of the stress response itself. In the absence of GCs, stress responses – high blood pressure, increased heart rate, and fever – are unbridled and could well be fatal. Essentially, GCs return activated blood and other cells to their quiescent state so that they stop producing the cytokines, PGs, and other inflammatory agents that constitute the host defense mechanism.

Under normal physiological conditions, COX-2 is regulated, among others, by endogenous GCs.

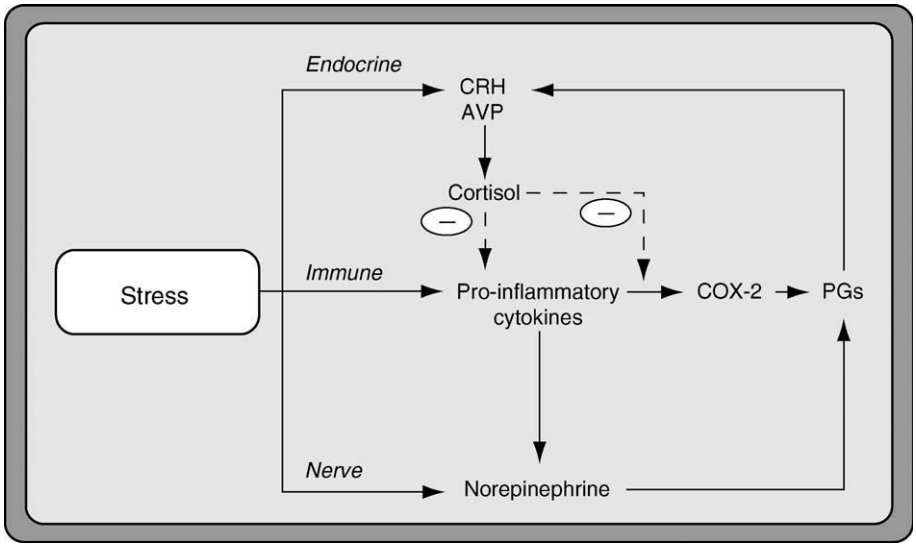


Figure 3 Participation and interdependence of endocrine, immune, and nervous systems in stress. AVP, arginine vasopressin; COX-2, cyclo-oxygenase-2; CRH, corticotropin releasing hormone; PGs, prostaglandins.

Whether the GC surge during stress – a pathophysiological state – affects concomitant PG synthesis has been examined using the rather crude technique of adrenalectomy (ADX) and, more recently, with GC-receptor antagonists such as RU486 (mifepristone). GCs do indeed moderate stress-induced PG production because PG synthesis is elevated in the absence of viable GCs and is reduced again by exogenous dexamethasone, but not by EPI, which is also lacking after ADX.

In the brain, suppression by GCs of PGE_2 along the HPA axis presents an interesting negative feedback response because at least one of the noradrenergic neural circuits leading to ACTH release is mediated via PGE_2. This PGE_2 is stimulated by circulating IL-1, and pro-inflammatory cytokines released into the plasma are also suppressed by GCs. Thus, GCs prevent their own stimulation. The inhibition of IL-1- and PGE_2-induced GCs by GCs could be part of a mechanism for the negative feedback between ACTH and GCs.

In general, and left to themselves, COX-2-stimulated PGs are injurious (they support inflammation allergy). The repression of COX-2 puts GCs into their familiar role of damage control by holding the body's stress response in check and trying to restore homeostasis.

See Also the Following Articles

Congenital Adrenal Hyperplasia (CAH); Caregivers, Stress and; Corticosteroid-Binding Globulin (Transcortin); Corticosteroids and Stress; Cushing's Syndrome, Neuropsychiatric Aspects; Glucocorticoids – Adverse Effects on the Nervous System; Holocaust, Stress Effects of; Hypotension, Hypovolemia, and Septic Shock.

Further Reading

Bernadini, R., Chiarenza, A., Calogero, A. E., et al. (1989). Arachidonic acid metabolites modulate rat hypothalamic corticotropin-releasing hormone secretion. *Neuroendocrinology* **50**, 708–715.

Brody, T. M., Larner, J. and Minneman, K. P. (1998). *Human pharmacology: molecular to clinical* (3rd edn.). St. Louis, MO: Mosby-Year Book.

Bugajski, J., Olowska, A., Gadek-Michalska, A., et al. (1996). Effect of indomethacin on the CRH- and VP-induced pituitary-adrenocortical response during social stress. *Life Sciences* **58**, 67–72.

Cowell, A. M. and Buckingham, J. C. (1989). Eicosanoids and the hypothalamic-pituitary axis. *Prostaglandins, Leukotrienes and Essential Fatty Acids* **36**, 235–250.

Elmquist, J. K., Scammell, T. E. and Saper, C. B. (1997). Mechanisms of CNS response to systemic immune challenge: the febrile response. *Trends in Neuroscience* **20**, 565–570.

Gether, U. (2000). Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. *Endocrine Review* **21**, 90–113.

Gilman, A. G., Rall, T. W., Nies, A. S. and Taylor, P. (1990). *The pharmacological basis of therapeutics* (8th edn.). New York: Pergamon Press.

Goppelt-Struebe, M. (1997). Molecular mechanisms involved in the regulation of prostaglandin biosynthesis by glucocorticoids. *Biochemical Pharmacology* **53**, 1389–1395.

Graham, D. J., Campen, D., Hui, R., et al. (2005). Risk of acute myocardial infarction and sudden cardiac death in patients treated with cyclooxygenase 2 selective and non-selective non-steroidal anti-inflammtory drugs: nested case-control study. *Lancet* **365**, 475–481.

Hata, A. N. and Breyer, R. M. (2004). Pharmacology and signaling of prostaglandin receptors: multiple roles in

inflammation and immune modulation. *Pharmacology and Therapeutics* **103**, 147–166.
Jabbour, H. N. and Sales, K. J. (2004). Prostaglandin receptor signalling and function in human endometrial pathology. *Trends in Endocrinology and Metabolism* **15**, 398–404.
Morimoto, A., Watanabe, T., Morimoto, K., et al. (1991). Possible involvement of prostaglandins in psychological stress-induced responses in rats. *Journal of Physiology* (London) **443**, 421–429.
Narumiya, S., Sugimoto, Y. and Ushikubi, F. (1999). Prostanoid receptors: structures, properties, and functions. *Physiological Reviews* **79**, 1193–1226.
Nasushita, R., Watanobe, H. and Takebe, K. (1997). A comparative study of adrenocorticotropin-releasing activity of prostaglandins E1, E2, F2 alpha and D2 in the rat. *Prostaglandins, Leukotrienes and Essential Fatty Acids* **56**, 165–168.
Oates, J. (1982). The 1982 Nobel Prize in physiology or medicine. *Science* **218**, 765–768.
Rivier, C. (1995). Influence of immune signals on the hypothalamic-pituitary axis of the rodent. *Frontiers in Neuroendocrinology* **16**, 151–182.
Vane, J. R., Bakhle, Y. S. and Botting, R. M. (1998). Cyclooxygenases 1 and 2. *Annual Review of Pharmacology and Toxicology* **38**, 97–120.
Watanabe, T., Morimoto, A., Morimoto, K., et al. (1991). ACTH release induced in rats by norepinephrine is mediated by prostaglandin E2. *Journal of Physiology* (London) **443**, 431–439.
Watanobe, H., Sasaki, S. and Takebe, K. (1995). Role of prostaglandins E1, E2 and F2 alpha in the brain in interleukin 1 beta-induced adrenocorticotropin secretion in the rat. *Cytokine* **7**, 710–720.
Zor, U. (1994). *Lipid mediators in health and disease*. London: Freund.

Proteases in Prokaryotes and Eukaryotic Cell Organelles

E Conway de Macario and A J L Macario
New York State Department of Health and The University at Albany (SUNY), Albany, NY, USA

Published by Elsevier Inc.

Protein Homeostasis
The Archaeal Proteasome
Other Proteases in Archaea
Compartmentalized Proteases in Bacteria
Membrane-Associated Proteolytic Machines in Bacteria
Proteases in Cyanobacteria, Mitochondria, and Chloroplasts
ATP-Independent Proteases

Glossary

AAA+ proteins A large family of proteins with ATPase capacity that participate in many cellular processes: AAA stands for ATPases associated with various cellular activities.

Archaea One of the three main evolutionary lines of descent or phylogenetic domains. The archaeal organisms are prokaryotes.

Bacteria One of the three phylogenetic domains. This domain includes all prokaryotes except archaea.

Chloroplast An organelle present in plant cells and photosynthetic protists capable of using light as source of energy by means of green pigments (chlorophyll) to synthesize carbon compounds from carbon dioxide and water. It is evolutionarily related to cyanobacteria.

Clp From caseinolytic protease, designates a large family of related enzymes that are also related with the Lon proteases and are present in bacteria, archaea, and eukaryotic organelles.

Eucarya (eukarya) One of the three main phylogenetic domains. This domain includes all the eukaryotes.

Mitochondrion An organelle present in animal and plant cells capable of using oxygen to generate energy (aerobes); it synthesizes most of the cellular ATP via oxidative phosphorylation. It is evolutionarily related to the alpha-purple bacteria.

Organelle A component of eukaryotic cells that resides within the confines of the cell membrane, in the cytosol, but is separated from the cytosol (subcompartment). Organelles contain many components with specialized functions that define their roles, e.g., respiration for the mitochondrion and photosynthesis for the chloroplast. Some organelles are considered to be derived from bacteria that entered into the eukaryotic cell, or its precursor, early in evolution.

Protein Homeostasis

A normal cell has many groups of proteins with distinctive functions. The members of each group must be in their native, functional conformation and within their physiological range of concentration if the cell is to remain fully functional and healthy. Maintenance of all proteins within their physiological conformations and concentration ranges, i.e., protein homeostasis, is achieved by a variety of means. Among these means are those that aim at assisting the nascent polypeptides to fold correctly, namely to achieve a native conformation, or to refold when the mature protein is damaged (unfolded) by stressors such as a sudden increase in temperature (heat shock), or a change in pH or in salinity. Molecular chaperones, discussed in other parts of this Encyclopedia, play a central role in protein folding and refolding.

Another aspect of protein homeostasis is the elimination of damaged molecules that are not reparable by the chaperoning systems, or the degradation of normal molecules that have become unnecessary to the cell after they have carried out their specific functions. Enzymes called proteases, usually forming large multimolecular assemblies such as the proteasome, play a central role in protein degradation and elimination, as discussed in this chapter. Thus the chaperoning and protein-degrading systems are two complementary arms of the whole mechanism responsible for protein homeostasis.

The Archaeal Proteasome

In archaea, the proteasome seems to be the central, if not the exclusive, energy-dependent protease. In contrast, the other prokaryotes, the bacteria, have several of these proteases, such as Lon and Clp (caseinolytic protein; discussed later), and some bacterial species have a proteasome in addition. In fact, the proteasome of one of these bacterial species, *Mycobacterium tuberculosis*, has been shown to be crucial for its survival in the face of stress.

Archaea have a 20S proteasomal particle, with the same overall structure as the eukaryotic counterpart. This overall structure also resembles that of the multimeric chaperonin complex GroEL/S typical of bacteria but also found in some archaea and in eukaryotic cell organelles. All of these complexes are barrel-shaped and have sevenfold rotational symmetry of the subunits arranged around a central cavity, which is accessible only from the ends. Despite the morphological similarities between the archaeal 20S proteasome and the folding chaperonin machine GroEL/S, these complexes differ in several important aspects, reflecting their very different roles and mechanisms of action. One of the key differences is the size of the mouth of the chamber in its opened state: the mouth in GroEL/S is about four times bigger than that in the proteasome. This difference is very important, because the gate's diameter plays a crucial role in the selection of what goes into the chamber and what is kept outside. Simply on the basis of maximum gate size, we see that entry into the destruction chamber of the proteasome is a highly selective process, allowing the access of only linearized polypeptides brought near by a specific mechanism that includes the attachment of a defined tag to the polypeptide. All of these selection mechanisms operating together determine exactly which polypeptide from among the very many in the intracellular molecular soup will make its way toward destruction. Obviously, a number of mechanisms seem to have evolved in the direction of preserving intracellular structures, a fact that is manifested most evidently in the positioning of the enzymatic sites. These occur inside a chamber without any possibility of digesting surrounding molecules, except those that are allowed inside after careful screening, selection, and preparation.

The archaeal 20S proteasome is composed of 28 subunits, which are related in primary structure and can be classified into 14 families encompassing two superfamilies, alpha and beta. The alpha subunits oligomerize to form the two outer rings, one on each end of the central barrel, i.e., the proteasome core that contains the proteolytic chamber and is formed by two rings of seven beta subunits each. Thus, the similarities with the 20S particle from eukaryotes are evident.

Archaeal and bacterial 20S proteasomes have two to four different subunits (isoforms) in each ring, depending on the species. In contrast, all subunits differ from one another in the eukaryotic proteasome. It should be mentioned, however, that studies in progress are demonstrating more variants of archaeal proteasome subunits, and it is becoming apparent that some species have more isoforms than others. Furthermore, studies in archaea, bacteria, and eukaryotes suggest that certain species, or perhaps many species, have more than one type of 20S particle, each built differently, using different isoforms, and probably having distinct roles. This may be inferred from the fact that in eukaryotes each subunit uniquely occupies a defined position in the ring, suggesting strict control of function and great specialization of the overall structure. Thus, the discovery of distinct types of 20S particles should not be surprising, considering the array of roles that the proteasome plays in a variety of cellular processes.

The archaeal proteasome does not have a 19S particle, in contrast to the eukaryotic 26S proteasome, but it does have a regulatory cap. The archaeal proteasome regulatory factors are called proteasome-activating nucleotidases (Pan), and they are gathered at the entrance of the 20S particle in a manner reminiscent of the 19S particle structure in eukaryotes. The eukaryotic 19S particle has two sections, a lid and a base, the former composed of eight subunits lacking ATPase capacity. These subunits are termed regulatory particle non-ATPases (Rpn). The base has at least six ATPase subunits, which belong to the AAA+, type I proteins, and are called regulatory particle triple A proteins (Rtp). In addition, the base has other subunits that are of the Rpn type. The archaeal regulatory structure, which sits on the alpha rings, is built of Rpt homologs, termed Pan.

AAA+ proteins are proteins involved in many cellular processes, which justifies their name: ATPases associated with various cellular activities, or AAA+.

The final steps of proteolysis, after the proteasome has concluded its work, involve the giant protease TRI (tricorn), which digests the oligopeptides emerging from the proteasome to tri- and dipeptides. These are finally cut into single amino acids by the TRI-interacting factors (peptidases) F1, F2, and F3.

The proteasome in archaea may cooperate with membrane-associated proteases, such as the Lon protease, to translocate and degrade membrane-associated proteins.

Other Proteases in Archaea

Archaea, like bacteria and eukaryotic organelles, possess Clp- and Lon-related proteases. The proteolytic systems of archaea display features that combine bacterial and eukaryotic elements. This duality is just another example of the many structures and functions found in archaea, which seem to be a mixture of bacterial and eukaryotic components, such as those observed in the chaperoning systems and in the gene regulatory machinery.

Compartmentalized Proteases in Bacteria

There are at least five ATP-dependent compartmentalized proteases in bacteria. These bacterial proteases have counterparts in eukaryotic organelles, e.g., mitochondria and chloroplasts. The role of these proteases is to maintain the pool of cellular proteins, such that each species of proteins is present within its physiological range of concentration, and all molecules are in their native, functional conformations. Proteases contribute to this protein homeostasis by eliminating protein molecules that are abnormal or those that are normal but have become unnecessary, dispensable, after they have done their work and that may interfere with the normal functions of the cell if allowed to stay around. Hence, these proteases are capable of demolishing tertiary and quaternary structures (e.g., they can disassemble multimeric complexes) and degrading their components.

Structure

The active protease structure (i.e., the proteolytic machine that recognizes the client polypeptides marked for degradation, prepares them and delivers them to the proteolytic sites, digests them, and releases the degradation products into the cytoplasm) is a large, barrel-like multimolecular complex. Like the proteasome, the basic bacterial protease structure is composed of multimeric rings stacked coaxially to enclose a central, cylindrical chamber, which is open at both ends. Typically, there are four rings; the central two rings contain the proteolytic sites on their inner surface. The peripheral rings, one on each end of the central proteolytic chamber, serve as substrate selectors and also capture the selected client polypeptide, prepare it (e.g., unfold it completely), thread the linearized substrate through the opening of the central cavity, and translocate it into the cavity. All of these activities, performed by the peripheral rings, require energy from ATP; thus, the components of these rings are ATPases. They belong to the large family of AAA+ ATPases.

Subunits

Several subunits are known to constitute the bacterial protease machines: ClpA, ClpB, ClpC, ClpE, ClpP, ClpX, and ClpY(HslU). Three of the main compartmentalized proteases in bacteria are ClpAP, ClpXP, and ClpXQ (or HslUY); these are composed of a protease (ClpP or ClpQ(HslV)) flanked by an ATPase ring formed by ClpA, ClpX, or ClpY(HslU), respectively. These flanking structures are formed by subunits, which are AAA + ATPases and which function in conjunction with the central structure (ClpP or ClpQ). However, they can also disaggregate protein aggregates by themselves.

Some of the genes encoding Clp proteins are stress inducible, for instance, *clpE*, *clpC*, and *clpP* in *Bacillus subtilis*; hence these proteins are members of the large group of stress proteins.

Rings

The subunits, six or seven of them together, associate to form rings with axial symmetry. Four of these rings

stack on top of each other to build the final structure. Since the rings associate coaxially, the final structure resembles a barrel with a chamber, open at both ends.

All of the subunits in each given ring are identical. The central two rings, which delimit the proteolytic chamber, are each formed by seven ClpP subunits. The two peripheral rings are formed by six subunits, either ClpA or ClpX (for ClpAP and ClpXP, respectively), and they sit on the central two-ring core, one on each end of this core. The peripheral rings in each final structure are identical, either ClpA or ClpX rings.

ClpB forms six-membered rings and interacts with the molecular chaperone machine (Hsp70(DnaK)-Hsp40(DnaJ)-GrpE) to form the Clp/KJE complex. This complex is a disaggregating machine capable of dissolving protein aggregates, particularly when these aggregates contain sHsp in addition to denatured polypeptides, which are the initial cause and key components of the aggregates.

ClpC is a stress protein (i.e., its parent gene is induced by stressors) that is apparently necessary for the survival of *B. subtilis* when this bacterium is faced with stressful conditions. ClpC forms complexes with ClpP, and the adaptor protein MecA (discussed later).

The ATPase peripheral rings play important roles when assembled together with the two central protease rings. These roles are substrate selection, capture, unfolding, and threading, and then translocation and delivery to the proteolytic chamber.

Substrate Selection

Proper selection of which protein will be disassembled and degraded is crucial. In the absence of selection, the entire pool of intracellular proteins (millions of them) would be destroyed rapidly and indiscriminately by the many protease machines extant inside the cell at all times. To avoid this chaotic protein destruction, there are essentially two mechanisms: (1) the enzymatically active sites are separated from the surrounding protein soup by the wall of the cylinder and (2) the peptides to be eliminated must have features that distinguish them from all the others and that make them recognizable by the peripheral rings as client polypeptides.

The proteasome of eukaryotes interacts with the ubiquitin system, which is the system for tagging the polypeptides that will then be recognized by the regulatory particle of the 26S proteasome. Bacteria do not have the ubiquitin system, but they do have tagging mechanisms that work under specific circumstances. In addition, the bacterial proteases can directly recognize partially denatured, damaged polypeptides through signals consisting of stretches of unstructured sequence, usually near one of the polypeptide's ends. Hydrophobic patches, exposed in these unstructured-sequence segments, are the key as signaling elements.

Also, some client polypeptides are tagged by bacterial tagging mechanisms, such as the SsrA tag system. This mechanism operates when defective polypeptides are produced by stalling ribosomes. In this situation, a short signal of 11 amino acids is added to the C-terminal end of the incomplete polypeptide. SsrA targets the client polypeptide for degradation by ClpXP, ClpAP, and other proteases; it is not specific to any one of them.

Adaptors

Adaptors are proteins that activate proteolytic machines, or that facilitate proteolysis by these machines, by bringing together client polypeptide and machine. The SsbB adaptor joins SsrA-tagged polypeptides together with ClpX (an ATPase) and thus enhances translocation of the client polypeptide into the ClpP proteolytic chamber. Other adaptors are Mub and MecA; the latter interacts with ClpC. MuB inhibits degradation or disassembly of certain potential client polypeptides by occluding the degradation signals present in their sequences.

Another adaptor is ClpS, which interacts with ClpAP to form the ClpAPS complex, a very competent and active protease.

Membrane-Associated Proteolytic Machines in Bacteria

In addition to the protease complexes discussed previously, all of which are in the cytoplasm, bacteria have other protease machines that are associated with their membranes. These machines carry out the degradation of membrane-associated client polypeptides. One of these machines is FtsH(HflB), which forms a very large complex (holoenzyme) with HflKC. FtsH also cooperates with the membrane-integrated metalloprotease HtpX for processing of metal-containing proteins.

Proteases in Cyanobacteria, Mitochondria, and Chloroplasts

In the cyanobacterium *Synechococcus elongatus*, several proteases, Clp-1, Clp-2, Clp-3, and ClpR, and the adaptor ClpS, have been identified. ClpS interacts with ClpA to form ClpAPS, as in other bacteria. Related proteases are found in the chloroplast, presumably evolutionarily related to those found in cyanobacteria.

Mitochondria, which are considered to be derived from ancient alpha purple bacteria – i.e., the ancestors of today's alpha purple bacteria – possess several proteases, including members of the two main AAA+ATPase families. These mitochondrial ATPases are associated with the inner membrane. In the mitochondrial matrix, there are homologs of the bacterial ClpB, ClpX, and Lon, named mtHsp78, mtMcx1, and mLon, respectively. Like their bacterial counterparts, these mitochondrial proteases form a proteolytic chamber with two stacked rings each made of seven identical subunits, and two regulatory rings, one at each end of the core proteolytic chamber, composed of six identical rings, which are distinct from those in the central core rings.

ATP-Independent Proteases

In addition to the ATP-dependent proteases discussed previously, there are others that do not require ATP. For example, one such group of proteases is the DegP (HtrA) family of serine proteases, which encompasses nearly 2000 members. They are widespread and are highly conserved in sequence, being present in the periplasmic space of Gram-negative bacteria and in the ER and chloroplast of eukaryotes.

Interestingly, DegP in the bacterium *Escherichia coli* has either chaperoning or proteolytic ability, depending on the temperature. At physiological temperatures, DegP functions as a chaperone, but at high temperatures it becomes a protease. Like the ATP-dependent proteases, DegP forms oligomeric rings (three subunits per ring) that stack on each other, coaxially, and thus form pairs of hexameric complexes enclosing a central cavity.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Heat Shock Genes, Human; Heat Resistance; Heat Shock Response, Overview; Viral Virulence and Stress; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol.

Further Reading

Borissenko, L. and Groll, M. (2005). Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes. *Journal of Molecular Biology* **346**, 1207–1219.

Dura, M. A., Receveur-Brechot, V., Andrieu, J.-P., et al. (2005). Characterization of a TET-like aminopeptidase complex from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Biochemistry* **44**, 3477–3486.

Franzetti, B., Schoehn, G., Hernandez, J.-F., et al. (2002). Tetrahedral aminopeptidase: A novel large protease complex from Archaea. *EMBO Journal* **21**, 2132–2138.

Guagliardi, A., Mancusi, L. and Rossi, M. (2004). Reversion of protein aggregation mediated by Sso7d in cell extracts of *Sulfolobus solfataricus*. *Biochemistry Journal* **381**, 249–255.

Mogk, A., Deuerling, E., Vorderwuelbecke, S., Vierling, E. and Bukau, B. (2003). Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Molecular Microbiology* **50**, 585–595.

Maupin-Furlow, J. A., Gil, M. A., Karadzic, I. M., Kirkland, P. A. and Reuter, C. J. (2004). Proteasomes: perspectives from the Archaea. *Frontiers in Bioscience* **9**, 1743–1758 [online journal].

Maurizi, M. R. and Xia, D. (2004). Protein binding and disruption by Clp/Hsp100 chaperones. *Structure* **12**, 175–183.

Pallen, M. J., Lam, A. C. and Loman, M. (2001). Tricorn-like proteases in bacteria. *Trends in Microbiology* **9**, 518–821.

Piszczek, G., Rozycki, J., Singh, S. K., Ginsburg, A. and Maurizi, M. R. (2005). The molecular chaperone, ClpA, has a single high affinity peptide binding site per hexamer. *Journal of Biological Chemistry* **280**, 12221–12230.

Schlee, S., Beinker, P., Akhrymuk, A. and Reinstein, J. (2004). A chaperone network for the resolution of protein aggregates: direct interaction of ClpB and DnaK. *Journal of Biological Chemistry* **336**, 275–285.

Tamura, N., Lottspeich, F., Baumeister, W. and Tamura, T. (1998). The role of Tricorn protease and its aminopeptidase-interacting factors in cellular protein degradation. *Cell* **95**, 637–648.

Zhang, X., Beuron, F. and Freemont, P. S. (2002). Machinery of protein folding and unfolding. *Current Opinion in Structural Biology* **12**, 231–238.

Proteases in the Eukaryotic Cell Cytosol

E Conway de Macario and A J L Macario
New York State Department of Health and
The University at Albany (SUNY), Albany,
NY, USA

Published by Elsevier Inc.

Introduction
Protein-Folding and Protein-Degrading Multimolecular Machines
The Protein Degradation Pathway Involving the Ubiquitin-Proteasome System (UPS)
The Proteasome
Steps Toward Degradation by the Proteasome
The Whole Picture

Glossary

ATP-dependent protease	A protease that requires energy from ATP hydrolysis to perform its functions.
Client polypeptide	A protein destined for degradation, proteolysis (the protease substrate).
Cofactors	Molecules (usually proteins) that cooperate with chaperones in directing the client polypeptide to the protease, and in preparing it (e.g., unfolding it) so that the protease can receive it for digestion.
Protease	An enzyme that digests proteins, breaking them down into components of various lengths and ultimately producing, in many cases, single amino acids. Different proteases differ in complexity: some are formed by a single molecule, whereas others are assemblages of several molecules.
Proteasome	An example of a proteolytic machine, typically present in the cytosol of eukaryotic and archaeal cells, and also in a small number of bacterial species.
Proteolytic machine	A protease that is a multimolecular assembly, in which each component plays a role in the series of events leading to protein digestion and release of cleavage products.
Self-compartmentalized protease	A proteolytic machine that has its active proteolytic sites inside a cavity, inaccessible to surrounding molecules, and thus requiring an active mechanism for entry of the substrate into the cavity for digestion.
Ubiquitin-proteasome system (UPS)	A protein-degrading system involving the proteasome and a series of enzymes. These enzymes are the ubiquitins (or other molecules with ubiquitin-like properties) that tag and prepare the client polypeptide for digestion in the proteasome.

Introduction

Small vs. Large Proteases

Protein homeostasis requires production of polypeptides as well as their regulated destruction, when they are no longer needed or are defective. Molecular chaperones play key roles in both the production of functional proteins and the destruction of useless ones. However, the central players in the destruction phase are the proteases. These enzymes occur, like the chaperones, in every cell and cell compartment in all organisms.

Proteases encompass a wide range of structures, from monomeric hydrolases to multimolecular, giant machines with molecular masses on the order of 1 million Daltons. The large molecular assemblies are self-compartmentalized and have their proteolytic sites inside a central chamber that can be reached only by selected substrates.

This article deals with the large, self-compartmentalized, multimolecular proteolytic machines in the three domains of life: archaea, bacteria, and eucarya (or eukarya).

A General Outline of the Process of Proteolysis

Proteins are, in general, degraded via a series of steps: (1) tagging, i.e., the polypeptide destined for degradation must have a tag, which makes it recognizable by the proteases or by the molecules that will direct the polypeptide to the protease; (2) unfolding, which is usually done via an ATP-dependent mechanism by specific components of proteases, such as Lon/Clp in bacteria and the 26S proteasome in eukaryotes (since at this stage the main event is the unfolding of the polypeptide, the molecules that do the unfolding are called unfoldases); (3) ATP-dependent proteases perform the initial enzymatic attack, generating oligopeptides with fewer than 30 amino acids; (4) these oligopeptides are then further digested by endopeptidases, tripeptidyl peptidases, and dipeptidyl peptidases, yielding still shorter oligopeptides (fewer than 10 amino acids long); and (5) these very short oligopeptides are finally digested into still shorter oligopeptides and single amino acids by aminopeptidases, carboxypeptidases, and di- and tripeptidases. Some of these proteases, e.g., TRI (tricorn) and TET (tetrahedral aminopeptidase), are large multimolecular

complexes like the proteasome. For example, TRI, which is present in some archaea and some bacteria (eukaryotes have in the cytosol a protease, tripeptidyl peptidase II (PPDII), which may be the functional equivalent of TRI), takes the oligopeptides produced by the proteasome and digests them to tri- and dipeptides, which are subsequently cleaved into single amino acids by peptidases called TRI-interacting factors F1, F2, and F3. In summary, unfoldases, the proteasome, TRI, and TRI-interacting factors constitute a multimolecular disassembly line. This general schema applies also to other protein-degrading complexes, some of which are discussed in this article.

Protein-Folding and Protein-Degrading Multimolecular Machines

Polypeptides misfolded because of a structural defect (mutation or posttranslational modification) or unfolded because of stress, as well as incompletely folded (nascent) normal polypeptides, in the cytosol confront a dangerous, highly competitive environment, mainly due to overcrowding. All of these kinds of polypeptides tend to entangle, stick to one another, and form aggregates and precipitates, which can be irreversible and toxic to the cell.

As the misfolded or unfolded polypeptides appear in the cytosol, they encounter the molecular chaperone machine formed by three major teammates, the proteins Hsp70(DnaK), Hsp40(DnaJ), and nucleotide exchange factor. The latter is GrpE (glucose-regulated protein E) in prokaryotes, and BAG-1 (BCL2-associated athanogene, where BCL2 stands for B cell lymphoma 2) or HspBP1 (Hsp70-binding protein 1) in eukaryotes. After this encounter, the polypeptide follows one of two alternative pathways, either toward folding or toward degradation, depending to some extent on its own characteristics and on which cofactors join the chaperone machine. In the eukaryotic cytosol, Hip (Hsc70-interacting protein) and Hop (Hsp70/Hsp90-organizing protein) promote folding, whereas BAG-1 and CHIP (C-terminus of Hsc70-interacting protein) direct the polypeptide toward degradation. In both cases, the polypeptide ends up inside a barrel-shaped chamber.

The folding and proteolytic machines are multimolecular assemblies with a central cavity or chamber. In the prokaryotes bacteria, the folding machine is the complex GroEL/S, and in the eukaryotic cytosol it is the CCT (chaperonin-containing TCP-1, also called TriC for TCP-1 ring complex, where TCP stands for tailless complex polypeptide). An equivalent of CCT, named thermosome, occurs in the prokaryotes archaea. In addition, some archaeal species also have GroEL/S-like bacteria.

The proteolytic machine in bacteria is the Clp complex, whereas in the eukaryotes and archaea it is the proteasome: 26S in the former and 20S in the latter. The organelles of the eukaryotic cell also have folding and proteolytic machines; for example, the mitochondrion has equivalents of the bacterial Clp complex for proteolysis, and it harbors equivalents of GroEL and GroES for protein folding.

The Protein Degradation Pathway Involving the Ubiquitin-Proteasome System (UPS)

In the eukaryotic cytosol, the pathway toward degradation includes various steps and involves diverse components. Initially, the molecular chaperone machine goes into action, and then the cofactors BAG-1 and CHIP interact with the machine and with the next group of participants in the pathway leading to the proteasome. This group is the ubiquitin system.

Ubiquitins

In order for a polypeptide that must be degraded to reach the proteasome, it must first be tagged so that the machine can recognize it. The tag is the ubiquitin molecule, and the tagging is done by the ubiquitin system, which consists of diverse enzymes sorted into three groups, E1, E2, and E3. The E1 are ubiquitin-activating enzymes that prepare ubiquitin for interaction with a carrier, the E2 ubiquitin-conjugating enzymes. E3 ubiquitin ligases bring together the polypeptide destined for degradation (substrate or client polypeptide) and E2, at which point the ubiquitin molecule is attached (ubiquitylation) to the client polypeptide. The proper tag that indicates proteolysis is composed of a chain of at least four ubiquitin molecules; peptides tagged with one to three ubiquitins are directed toward other cellular processes and not to degradation by the proteasome (although they may be directed toward degradation by the lysosome). It must be emphasized that ubiquitylation is widespread in the cell and participates in many processes distinct from protein degradation: cell cycle, DNA repair, regulation of gene transcription, apoptosis, and others.

Ubiquitylation

Ubiquitin is a 76-amino-acid polypeptide, and ubiquitylation occurs via formation of an isopeptide bond between an internal lysine of the substrate and the C-terminal glycine (glycine 76) of ubiquitin. Successive additions of more ubiquitins to form a polyubiquitin chain of at least four units, i.e., a good signal for the proteasome, use the same mechanism

and bond type, but the ubiquitin residues involved are glycine 76 and lysine 48: new ubiquitins are attached to an internal lysine of the last ubiquitin molecule in the chain. A polyubiquitylated substrate is then recognized by the proteasome.

Chaperones and Cofactors for Protein Degradation

A pivotal component of the chaperone machine, and thus of the folding and degradation pathway, is Hsp70. This molecule has segments, motifs or domains, that interact with the other components of the machine and, most importantly, with the client polypeptide in need of assistance for folding or refolding, and also with the polypeptide beyond recovery that must be degraded.

The client polypeptide usually displays hydrophobic residues that, in a folded, native protein, lie in the interior of the molecule, inaccessible to other molecules in the immediate surroundings. Incompletely folded, partially unfolded (denatured), and misfolded polypeptides expose hydrophobic patches on their surface, and these attract Hsp70, which has a peptide-binding domain. This domain is flanked, in the linearized Hsp70 sequence, by an ATPase domain on the N-terminal side and by a lid on the C-terminal side, including the sequence glutamate-glutamate-valine-aspartate (EEVD) at the C-terminus.

The N-terminal section of Hsp70 functions as an ATP-binding and ATPase domain; in addition, it is bound by a number of cochaperones or cofactors of the BAG family (e.g., BAG-1 through 6 and HspBP1) and by Hip. With its C-terminal domain, Hsp70 can bind CHIP and Hop.

BAG-1 functions as a nucleotide exchange factor. The various BAG proteins each have on their C-terminal halves the BAG domain that enables them to bind Hsp70 as well as several other proteins (since they participate in a number of cellular processes). In addition to the BAG domain situated near the C-terminus, BAG family members have a ubiquitin-like domain in the middle of the molecule. This domain can bind to the proteasome and can thus direct a client polypeptide already bound to Hsp70 toward degradation.

CHIP has a tetratricopeptide repeat (TPR) domain with several copies of the TPR in its terminal half, a U box near its C-terminus, and a coiled-coil motif in between. This latter motif participates in protein–protein interactions, whereas the TPR domain is involved in the interaction of CHIP with the C-terminus of Hsp70, and the U box enables CHIP to act as an E3 ubiquitin ligase. The activity of CHIP mimics that of an E3 ligase: it binds to the Hsp70/substrate complex and then the E2 molecule Ubc4/5, and it thus facilitates the enzymatic polyubiquitylation of the substrate with the participation of an E2 enzyme.

A TPR motif or domain consists of a series of TPRs in tandem; the TPRs are imperfect repeats of 34 amino acids each. This motif is present in many proteins and participates in a great variety of molecular interactions.

In summary, both BAG-1 and CHIP bind the C-terminus of Hsp70: BAG-1 via its C-terminal BAG domain and CHIP via its N-terminal TPR domain. CHIP directs the client polypeptide to the proteasome in association with Ubc4/5 or in association with Ubc4/5 and BAG-1. Another chaperone, Hsp90, comes into play for other similar processes involving specific substrates, e.g., steroid hormone receptors. Here, the general mechanism is essentially the same as the one just described for those cases in which Hsp90 does not participate.

The Proteasome

The proteasome is one of the large protease complexes that have a central chamber, roughly cylindrical, in which proteolysis takes place. Other examples of similar proteolytic chambers are the Clp complexes that are found in the cytoplasm of bacteria and inside organelles in eukaryotes.

The proteasome has been found in the cytosol and nucleus of eukaryotic cells, in the cytoplasm of archaea, and in the cytoplasm of some bacteria, such as gram-positive actinomycetes (e.g., *Mycobacterium tuberculosis*). The eukaryotic proteasome is identified as the 26S particle, whereas those of archaea and bacteria are identified as 20S particles, according to their approximate sedimentation coefficients.

Structure

The 26S proteasome is composed of two main particles (or subparticles), the 20S and the 19S particles, both of which are polymers typically composed of 28 and 17–19 subunits, respectively. The 20S is the proteolytic particle, whereas the 19S particle is regulatory. In current terminology, 26S particle denotes a 20S particle with one 19S particle on one end and a 20S particle with two 19S particles, one on each end.

The 20S particle is formed by four stacked rings, two alpha and two beta rings, the latter in the middle, and one alpha ring on each end of the central barrel formed by the central beta rings. Each ring in turn is constituted of seven subunits, alpha and beta for the alpha and beta rings, respectively. All of the alpha subunits and all of the beta subunits are related in primary structure, but not identical. These subunits are arranged in rings with a central cavity, and since

the four rings are stacked coaxially in the overall structure, the 20S proteasome can be described as a cylinder, resembling a barrel.

The 19S particle is composed of at least 17 distinct subunits arranged in two complexes (or subcomplexes), the base and the lid. The base subcomplex sits on the alpha ring of the 20S particle and is formed by nine subunits, of which six are ATPases. The lid is distal to the base subcomplex, in contact with the intracellular milieu, and is formed by eight subunits devoid of ATPase capacity.

Role of the Proteasome Subunits and Particles

While the 20S particle is involved in the actual degradation (proteolysis) of the client polypeptide, the 19S particle has other functions, such as (1) recognition and sorting of the client polypeptides, i.e., those correctly ubiquitylated with chains of four or more ubiquitins; (2) opening of the mouth (gate) of the proteasome to allow entry of the substrate; (3) unfolding of the client polypeptide as much as necessary for it to negotiate the entrance to the proteasome, which is very narrow and does not allow passage of folded proteins; and (4) threading of the extended, unfolded substrate through the gate and transfer of it to the alpha ring of the 20S particle and beyond, so that the polypeptide reaches the proteolytic chamber delimited by the central beta rings. All of these steps require energy from ATP hydrolysis, which is why the base subcomplex has six ATPase subunits.

Steps Toward Degradation by the Proteasome

Calling for Destruction

The polypeptide destined for degradation has short stretches (5–25 amino acids) that constitute hydrophobic patches and mark it for destruction. However, the proteasome does not recognize such signals; for this to happen, a tag, the polyubiquitin chain described earlier, must be added to the client polypeptides. There may be exceptions to this rule. It is possible that some E3 enzymes can interact directly with the proteasome or that some client polypeptides can be recognized directly by the proteasome and engulfed for degradation without participation of E1 and E2 enzymes. These would be examples of what is called ubiquitin-independent protein degradation by the proteasome.

Unfolding, Release of Chaperones, Gate Opening, and Threading

The client polypeptide cannot enter the mouth of the proteasome if it is not opened first and if the polypeptide is not linearized and freed from the chaperones and cochaperones accompanying it before it is offered to the mouth. Thus, a mechanism for unfolding polypeptides operates at the mouth of the proteasome. Chaperone–cochaperone release, polypeptide unfolding, and gate opening are probably functions of the 19S particle, but the mechanism has not yet been elucidated.

Translocation from the Cytosol into the Chamber

The unfolding, threading, and translocation of the client polypeptide are most likely simultaneous events, although unfolding may partially precede threading, which in turn may begin before the start of translocation and proceed during translocation, and the final moments of the translocation may occur after the preceding events have concluded. By this time, the leading end of the polypeptide will have reached the proteolytic chamber, with proteolysis starting immediately. It is possible that even proteolysis begins almost simultaneously with translocation.

Proteolysis

Unstructured, less stable segments of the incoming, extended polypeptide are probably recognized by the proteolytic sites lining the inside of the proteolytic chamber, and digestion starts immediately. This part of the process is not specific. Other steps, preceding the actual digestion inside the proteasome, are selective: they depend on specific interactions between molecules that possess certain markers that make them interact only in certain combinations and, thus, recognize and process substrates with signals that are specific for only one of these combinations. In these steps, the role of the E enzymes is crucial, as is explained later.

Indeed, the proteolytic systems are finetuned so as to avoid indiscriminate digestion of cellular proteins and to steer toward destruction only those that must be eliminated for the good of the cell. The entire protein degradation process is strictly regulated by a number of mechanisms involving the E enzymes and the 19S proteasome components, as well as other mechanisms that insert specific signals on those polypeptides that are not wanted in the cell.

The proteolytic efficiency inside the proteasome appears to be very high, possibly a consequence of a concentration of enzymatic sites and the ease of encounter and contact between these sites and the extended substrate sliding naked, as it were, in close proximity of the sites. The resulting products are oligopeptides of 3–30 amino acids.

Exit of Digestion Products

Oligopeptide exit from the proteasome seems to be regulated, too, perhaps as strictly as the selection and entrance of the client polypeptide. However, little is known of the exit process. We can assume that rapid escape must be avoided, so that there will be no generation of incompletely digested substrate fragments, i.e., incomplete degradation. Likewise, excessive retention of completely digested fragments would have to be avoided lest the proteolytic chamber and the exit gate be clogged; such clogging would slow the flux of the whole process, including the pace of the proteolytic activity.

Processing and Reuse of Digestion Products

The oligopeptides emerging from the proteasome are cleaved into shorter oligopeptides and finally into single amino acids by aminopeptidases, carboxypeptidases, and di- and tripeptidases, e.g., in archaea TRI and TET, which are large multimolecular complexes like the proteasome, and TRI-related proteases in eukaryotes. The amino acids generated are used again in the synthesis of new polypeptides on the ribosome.

The Immunoproteasome

The proteasome also plays a role in immunity: it processes antigen and produces antigenic oligopeptides for display on the cell surface associated with other molecules as required for triggering a certain group of T lymphocytes. This type of antigen processing is the function of a specialized proteasome, the immunoproteasome, which differs in some details from the other proteasome discussed earlier, which is involved in general protein degradation.

Proteasomes at inflammatory sites undergo induction of regulatory subunits in the 19S particle and changes in the proteolytic subunits in the 20S particle: subunits beta1, beta2, and beta5 are replaced by their immuno counterparts, beta1i (LPM2), beta2i (MECL-1), and beta5i (LMP7), respectively. The 20S particle is assembled *de novo* with the immuno and the regular subunits in, for example, cells stimulated by γ-interferon or tumor necrosis factor. The other, so-called constitutive proteasome is present in normal cells and in immunoprivileged locales, e.g., brain tissue.

Proteins that appear as antigens to the organism are degraded in the immunoproteasome, and the resulting oligopeptides (8–10 amino acids long) are transported from the cytosol to the endoplasmic reticulum (ER) with the participation of TAP (transporter associated with antigen processing) proteins. In the ER, the peptides bind MHC (major histocompatibility complex) class I molecules, and the complex is inserted on the cell surface. This type of antigen presentation is critical for the action of certain T lymphocytes, e.g., cytotoxic T cells recognize and interact with the peptide–MHC complex, which may lead to the lysis of the cells carrying the complex on the surface as a means of eliminating potentially damaging cells.

It is believed that the immunoproteasome generates correct epitopes from the C-terminus of the protein antigen, but other enzymes are needed for generation of N-terminus epitopes. Some of the latter are LAP (leucine aminopeptidase) and ERAP 1 (ER aminopeptidase 1), which, as its name suggests, is associated with the ER.

Proteasome Pathology

The proteasome, and the immunoproteasome, can be defective and can thus compromise cell physiology and health. Since the proteasome is the major protagonist in the elimination of proteins that are not beneficial to the cell, when the proteasome function is impaired or stops, the effects are immediate and considerable. Likewise, failure of the immunoproteasome will have an impact on several functions of the immune system, with repercussions throughout the entire body.

Proteasome inhibition by chemical compounds, genetic defects in the component subunits, and acquired alterations of these subunits have all been shown to result in proteasome malfunction and in pathological accumulation of protein aggregates. This effect can be even more deleterious in proteinopathies (e.g., Alzheimer's, Parkinson's, and Huntington's disease). The pathogenesis of these diseases includes misfolding and aggregation of genetically defective proteins, which form intra- and extracellular precipitates in various tissues, particularly the brain, resulting in cell death. Proteasome failure can only aggravate this pathogenetic sequence. Also, it has been observed that the proteasome suffers alterations as the process of senescence progresses.

The Whole Picture

The diversity of the proteasome is evidenced by the variety of subunits encoded in the genomes already sequenced. Likewise, the conservation of the proteasome is apparent in the similarity among all of the subunits, particularly within a single type, alpha or beta, and in the overall architecture, which essentially consists of stacked rings of subunits assembled together to enclose a central cavity, with axial gates at both ends, in all species known to have proteasome.

In eukaryotes, proteasomes have been found in the cytosol and nucleus, and in association with the ER

and the cytoskeleton. Also, proteasomes have been found in archaea and in some bacteria, as mentioned earlier.

It has been demonstrated, or it is believed based on some experimental data, that the proteasome plays a key role not only in protein quality control as discussed in this article, but also in several other cellular processes, such as the immune response, cell division, metabolism, DNA repair, and control of gene transcription. In these instances, the ubiquitin system also intervenes.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Chaperonopathies; Heat Shock Genes, Human; Heat Resistance; Heat Shock Response, Overview; Proteases in Prokaryotes and Eukaryotic Cell Organelles; Viral Virulence and Stress.

Further Reading

Ciechanover, A. (2005). Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nature Reviews Molecular Cell Biology* **6**, 79–86.

Ciechanover, A. and Iwai, K. (2004). The ubiquitin system: from basic mechanisms to the patient bed. *IUBMB Life* **56**, 193–201.

Esser, C., Alberti, S. and Hoehfeld, J. (2004). Cooperation of molecular chaperones with the ubiquitin/proteasome system. *Biochimica et Biophysica Acta* **1695**, 171–188.

Geier, E., Pfeifer, G., Wilm, M., et al. (1999). A giant protease with potential to substitute for some functions of the proteasome. *Science* **283**, 978–981.

Glickman, M. H. and Ciechanover, A. (2002). The ubiquitin-proteasome proteolytic pathway: Destruction for the sake of construction. *Physiology Reviews* **82**, 373–428.

Goldberg, A. L. (2003). Protein degradation and protection against misfolded or damaged proteins. *Nature* **426**, 895–899.

Kisselev, A. F., Akopian, T. N., Woo, K. M. and Goldberg, A. L. (1999). The sizes of peptides generated from protein by mammalian 26 and 20S proteasomes: implications for understanding degradative mechanisms and antigen presentation. *Journal of Biological Chemistry* **274**, 3363–3371.

Kloetzel, P. M. and Ossendorp, F. (2004). Proteasome and peptidase function in MHC-class-I-mediated antigen presentation. *Current Opinion in Immunology* **16**, 76–81.

Lum, R., Tkach, J. M., Vierling, E. and Glover, J. R. (2004). Evidence for an unfolding/threading mechanism for protein disaggregation by *Saccharomyces cerevisiae* Hsp104. *Journal of Biological Chemistry* **279**, 29139–29146.

Maupin-Furlow, J. A., Gil, M. A., Karadzic, I. M., Kirkland, P. A. and Reuter, C. J. (2004). Proteasomes: Perspectives from the Archaea. *Frontiers in Bioscience* **9**, 1743–1758.

Muratami, M., Kung, C., Shokat, K. M. and Tansey, W. P. (2005). The F box protein Dsg1/Mdm30 is a transcriptional coactivator that stimulates Gal4 turnover and cotranscriptional mRNA processing. *Cell* **120**, 887–899.

Ortega, J., Heymann, J. B., Kajava, A. V., Ustrell, V., Rechsteiner, M. and Steven, A. C. (2005). The axial channel of the 20 S proteasome opens upon binding of the PA200 activator. *Journal of Molecular Biology* **346**, 1221–1227.

Vaux, D. L. and Silke, J. (2005). IAPs, RINGs and ubiquitylation. *Nature Reviews Molecular Cell Biology* **6**, 287–297.

Zhang, X., Beuron, F. and Freemont, P. S. (2002). Machinery of protein folding and unfolding. *Current Opinion in Structural Biology* **12**, 231–238.

Protein Assembly *See:* Protein Synthesis; Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Response, Overview; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Protein Folding *See:* Protein Synthesis; Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Response, Overview; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Protein Synthesis

M A Brostrom and C O Brostrom
Robert Wood Johnson Medical School, Piscataway, NJ, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 273–282, © 2000, Elsevier Inc., with revisions made by the Editor.

Phases of Protein Synthesis
Regulation of Translational Initiation
Cellular Stress and the Inhibition of mRNA Translation
Comparable Roles of the Endoplasmic Reticulum and Cytoplasmic Stress Responses in Recovery of Translational Activity
elF2 Kinase as a Focal Point for Regulating Protein Synthesis during Stress

Glossary

Chaperone	Protein that functions to support other proteins or growing polypeptides in a particular conformation.
Cytoplasmic stress	Condition associated with the accumulation of misfolded or damaged proteins in the cytoplasm.
Endoplasmic reticulum (ER)	Subcellular organelle in which newly synthesized secretory, lysosomal, and membrane proteins are processed and folded.
ER stress	Condition associated with the accumulation of unfolded or underprocessed proteins in the ER.
GRP78/BiP	A glucose-regulated 78-kDa ER-resident chaperone induced in response to ER stress.
HRI	Heme-regulated elF2 kinase that functions to inhibit translational initiation in erythroid cells.
HSP/HSC70	A 70-kDa cytoplasmic chaperone(s) induced in response to cytoplasmic stress.
PKR	An interferon-inducible, double-stranded RNA-activated protein kinase that phosphorylates eukaryotic initiation factor (elF) 2 to inhibit translational initiation.
PKR-like ER kinase	An ER membrane-spanning elF2 kinase that phosphorylates elF2 in response to ER stress.
Proteotoxic	Condition resulting in the misfolding or damage of cellular proteins.
Translational elongation	Process occurring on 80S ribosomes through which amino acids are transferred from aminoacyl-tRNAs to the growing polypeptide chain sequentially in accord with the internal codons of mRNA.
Translational initiation	Process by which the 80S ribosome complex carrying the initiator tRNA, met-tRNA, and mRNA properly phased at the initiation codon is assembled.

Protein synthesis refers to the biological process whereby amino acids are assembled by peptide bonding into specific polypeptide sequences in accord with genetic blueprints encoded by deoxyribonucleic acid (DNA).

Phases of Protein Synthesis

The overall process of protein synthesis occurs in several phases, each of which is highly complex and generally viewed as a separate field of study.

DNA-Directed Nuclear Transcription of Temporary Ribonucleic Acid Constructs (Messenger RNA) and the Processing of mRNA into Ribonuclear Particles for Movement to the Cytoplasm

Information transfer from DNA to mRNA is reflected by the synthesis of analogous purine/pyrimidine ribonucleotide sequences arranged in triplet groupings (codons) that specify individual amino acids. General features of the mRNA structure include a 5′ to 3′ phosphodiester linkage of nucleotides, a 5′-methylated guanylate cap, a 5′-untranslated region, a coding region, and a 3′-polyadenylate tail. Nucleotide base pairing produces variable degrees of internal looping and winding. A number of stress-related hormones, including adrenal steroids and thyroxine, alter the synthesis of mRNAs encoding specific proteins. The relationship of these transcriptional alterations to the general slowing of protein synthesis produced by these hormones in many tissues is unclear.

Ribosomal Synthesis of Protein through Translation of mRNA Codons

This process is divisible into two phases: translational initiation, which relates to the loading of ribosomes onto mRNA, and translational elongation, which refers to the linkage of amino acids on the ribosomal surface by peptide bonding to produce a growing polypeptide. Eukaryotic translational initiation involves the binding of the small (40S) ribosomal subunit near the mRNA cap and its subsequent movement

(scanning) in a 5′–3′ direction until it detects the first AUG (initiator) base sequence. This triplet ordinarily provides the starting point for reading additional codons specifying the amino acid sequences for assembling proteins. Initiation is catalyzed by at least nine distinct proteins termed initiation factors (elF) composed of some 25 assorted protein subunits. A number of these factors are subject to alterations of catalytic activity through protein phosphorylation by various protein kinase activities. Scanning involves an intricate interplay of these initiation factors with the 40S ribosomal subunit and mRNA. The 40S ribosomal subunit is converted to a 43S preinitiation complex by association with elF3 and with a ternary complex consisting of elF2, methionyl initiator transfer RNA (met-$tRNA_i$), and GTP. Binding of elF4F (a complex of 4E, 4A, and 4G) and elF4B produces helicase mediated unwinding (melting) of the cap region of mRNA that permits cap-proximal association of the 43S preinitiation complex. Stabilization of this complex with subsequent scanning of the ribosomal subunit to the initiator AUG appears to require at least two additional factors, elF1A and elF1. Upon binding of the anticodon of met-$tRNA_i$ to the initiator AUG by codon–anticodon base pairing, elF5 stimulates hydrolysis of the GTP associated with elF2. This hydrolysis is associated with the binding of the initiator met-$tRNA_i$ at the 40S P- (peptidyl binding) site, the release of all initiation factors from the 40S ribosomal subunit, and the binding of the large (60S) ribosomal subunit to form an 80S ribosome competent of performing peptide chain elongation.

Translational elongation involves the addition of amino acids to the growing peptide chain as specified by the sequence of mRNA codons occurring after the initiator AUG. The process requires two elongation factors (eEF) and a series of transfer RNAs, each of which possesses a specific anticodon and is subject to enzymatic aminoacylation with a specific amino acid. Peptide chain assembly proceeds through eEFl-mediated binding of aminoacyl-tRNA molecules to a binding site (A site) on the ribosome such that the anticodon of the tRNA associates with the codon of the mRNA. The carboxyl end of the peptide chain (or the initiator methionine) of the tRNA in the P site is uncoupled from the tRNA and is covalently linked by peptide bonding to the amino group of the aminoacylated tRNA in the A site. Peptide bonding is catalyzed by a peptidyl transferase activity associated with the ribosomal RNA of the 60S subunit. In conjunction with the hydrolysis of GTP, eEF2 supports the translocation of the resultant peptidyl-tRNA from the A site to the P site and the movement of the ribosome three nucleotides (1 codon) along the mRNA molecule. The A site is then available for binding another aminoacyl-tRNA in accord with the specificity of the next codon on mRNA. The process of peptide bond formation and ribosomal movement continues until a stop codon (UAA, UAG, or UGA) is reached on the mRNA. In conjunction with a release factor, the peptidyl transferase catalyzes the hydrolysis of the peptidyl-tRNA linkage, permitting the release of the completed peptide.

Posttranslational Protein Processing

Newly synthesized polypeptides may be released directly into the cytosol or targeted for translocation into the endoplasmic reticulum (ER) by a short series of amino acids at the amino terminal end of the protein termed the signal peptide sequence. In either event the polypeptides are biologically inactive until folded to proteins with specific three-dimensional structures. The development of biological activity frequently requires the binding or covalent linkage of various prosthetic groups and the association of polypeptides into proteins consisting of multiple (multimeric) subunits. Cotranslational protein folding in the cytoplasm (cytosol plus mitochondria) is assisted by various protein chaperones, including a series of proteins termed heat shock proteins (HSPs) that are induced by conditions that damage cytosolic proteins. Protein folding associated with the ER occurs simultaneously with contranslational translocation and is assisted by various ER resident chaperones, including GRP78/BiP, GRP94, protein disulfide isomerase, calnexin, and others. The signal peptide is normally removed by an endopeptidase. Some proteins are retained in association with the ER membrane, frequently looped back and forth between the cytosolic and lumenal faces of the ER in up to sevenfold membrane-spanning units. These proteins are integral to the synthesis of new membrane. A large fraction of the protein processed within the ER lumen is destined for extracellular secretion. These proteins are frequently stabilized by various types of glycosylation, disulfide bond formation, the oxidation of specific proline residues, and assembly into multimeric proteins. Disulfide bond formation is supported by an ER redox potential that is relatively more oxidizing than that of the cytoplasm. Protein processing is also supported by the high Ca^{2+} content of the organelle. Proteins that meet certain folding criteria move by vesicular transport to the Golgi for additional modification, whereas misformed proteins are targeted for degradation. Inhibitors of ER protein processing that slow protein trafficking to the Golgi result in the induction of some chaperones, most notably GRP78 and GRP94.

Regulation of Translational Initiation

The regulation of translational initiation involves the selection frequency of specific mRNA transcripts for translation and alterations in the rates of ribosomal loading on the transcripts.

Frequency of Transcript Selection

Selection frequency governs the frequency at which individual proteins are synthesized in preference to other proteins. This parameter is affected by the relative abundance of each species of mRNA transcript, which in turn is a function of rates of mRNA transcript turnover, by the recruitment of transcripts from ribonuclear particles into active translation, and by the susceptibility of the transcript to cap-proximal melting by the elF4 complex, permitting loading of 43S preinitiation complexes. Interaction of the cap-binding protein, elF4E, is influenced by the 5′ leader structure of mRNA such that differential rates of utilization occur with different species of mRNA. The activities of various components of the elF4 complex, including 4E and 4G, are increased by protein phosphorylation occurring in response to mitogens and hormonal growth factors. Two factors, 4E-BP1 and 4E-BP2, have been reported to exist that modify the activity of elF4E. The dephosphorylated form of 4E-BP1 is an inhibitor of elF4E. Phosphorylation and dissociation of 4E-BP1 occur in response to growth promoters such as insulin and epidermal growth factor, whereas certain viral infections, including polio and encephalomyocarditis virus, result in dephosphorylation of this protein. A number of protein kinase activities participate in these phosphorylations in cell-free preparations, including protein kinases A and C, S6 ribosomal protein kinase, and certain kinases for which casein serves as a substrate. Continuing efforts are in progress to identify the endogenous protein kinase(s) that catalyzes these phosphorylations. Phosphorylation of either 4E or 4G is associated with increased rates of initiation. While ribosomal loading ordinarily proceeds through mRNA cap-dependent association, it can also occur through internal loading of ribosomes. Internal ribosomal loading is thought to occur on certain mRNA with unusually long 5′-untranslated regions, such as that for GRP78, various HSPs, and some virally derived messages. elF4E is proposed to promote internal ribosomal loading.

Alterations in Rates of Ribosomal Loading

Formation of the 435 preinitiation complex from the association of the 40S ribosomal subunit with the ternary complex of elF2, GTP, and met-tRNA, provides a second major regulatory point in translational initiation (Figure 1). The 435 complex subsequently associates with mRNA, scans to the initiator AUG, and binds the 60S ribosomal subunit to form a monosome capable of translational elongation. At this point the elF2-associated GTP is converted to GDP, and the factor dissociates from the ribosome. Because the binary elF2–GDP complex does not support the formation of the ternary complex with met-tRNA$_i$, a subsequent round of initiation requires the regeneration of elF2–GTP via the catalytic exchange of GDP for GTP. This exchange is accomplished by elF2B, a factor typically present at low stoichiometric ratios with respect to elF2. The affinity of elF2B for elF2 is increased greatly by the phosphorylation of Ser-51 of the α subunit of elF-2 by specific elF2 kinase activities described later. Increased phosphorylation of elF2α is recognized to mediate the translational repression occurring in mammalian cells subjected to a wide variety of physical, chemical, and nutritional stresses. A 20–30% increase in the phosphorylation of elF2α is frequently adequate for the sequestration of elF2B into an inactive complex such that recycling of elF2 is largely inhibited. Cells, however, tend to vary somewhat in their relative contents of elF2B and elF2 and in the degree of elF2α phosphorylation required for translational suppression. There is reason to suspect that elF2B may be subject to additional regulatory inputs. For example, the catalytic activity of the factor is thought to be increased by an insulin-mediated phosphorylation.

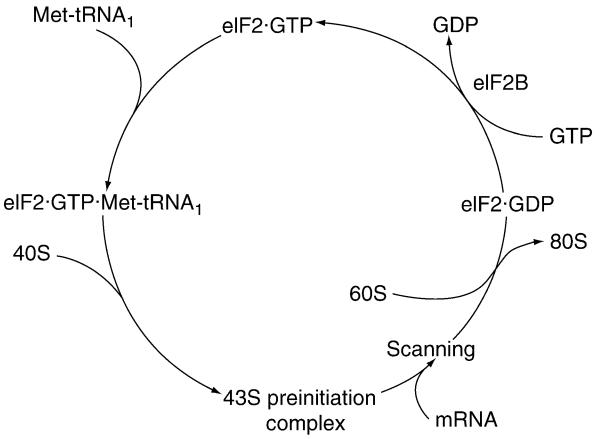


Figure 1 The elF2 cycle.

Phosphorylation of elF2 at Ser-51 is catalyzed by four well-characterized protein kinases termed GCN2, PKR, PERK, and HRI. These enzymes comprise a distinct class of protein kinases that share significant sequence homology in their catalytic domains. GCN2 appears to be expressed selectively in yeast and is activated by amino acid starvation. PKR is distributed broadly among mammalian cell

types, inducible by interferon, and activated by low concentrations of double-stranded RNA and by certain viral RNAs. PKR is thought, therefore, to mediate the antiviral effects of interferon. As described later, PKR appears to phosphorylate eIF2 during cytoplasmic stress. PERK functions as an ER membrane-spanning eIF2 kinase that mediates eIF2 phosphorylation during ER stress. The cytoplasmic domain of PERK possesses catalytic activity whereas its lumenal domain is thought to sense ER stress. HRI, the heme-regulated inhibitor kinase, is activated by heme deficiency and is expressed predominantly by erythroid cells. Both PKR and HRI undergo an autophosphorylation and dimerization during activation. While less well characterized, PERK probably undergoes similar modifications during activation. PKR, which resides on 60S ribosomes bound to the cytosolic face of the ER, may be capable of cross-dimerizing with PERK.

Most mRNA actively engaged in translation load multiple ribosomes and therefore sediment as polysomal assemblies during density gradient centrifugation. The rate of ribosomal loading in conjunction with the rate of peptide bond formation determines the overall rate of translation and the polyribosomal content of cells. Polyribosomal content and size tend to increase whenever the rate of peptide chain elongation is rate limiting with respect to initiation. While some additional recruitment of mRNA from ribonuclear particles into polyribosomes may occur through the phosphorylation and activation of the eIF4 complex in response to mitogens or growth factors, it is unclear that sufficient initiation factors and ribosomal subunits are generally available on a short-term basis to support large increases in initiation through this mechanism. Growth factors also have the potential to alter peptide chain elongation rates through the phosphorylation and activation of eEF1 and, perhaps, through the phosphorylation of ribosomal protein S6 or of various aminoacyl-tRNA synthases. The overall promotion of translation by these modifications as measured by amino acid incorporation ranges on the order of 30–100% depending on cell type over 1–2 h. In contrast, the phosphorylation of eIF2α in response to various conditions producing cellular stress inhibits initiation by approximately 85–95% within a few minutes in conjunction with reductions in polyribosomal contents. Under these conditions, provision of the 43S preinitiation complex, rather than recruitment of mRNA transcripts through cap melting considerations, becomes rate-limiting on initiation. Suppression of 43S complex formation appears to favor internal ribosomal loading on mRNA transcripts with long 5′ leaders typical of various HSPs and of GRP78. Overall, the eIF2 input into the regulation of translational initiation provides an adjustable braking system that overrides more slowly imposed stimulatory controls. Potential regulation emanating through protein phosphatase activities for the various initiation factors remains largely unexplored.

Cellular Stress and the Inhibition of mRNA Translation

Conditions that either damage intracellular proteins or produce an accumulation of protein-folding intermediates induce cellular stress responses, as defined by the increased expression of one or more HSP or GRP protein chaperones. Protein synthesis is suppressed at translational initiation in a manner that coincides with increased phosphorylation of eIF2α. Alterations in the content or the degree of phosphorylation of the eIF4 complex have also been observed following varying degrees of cellular damage. Current evidence favors a model in which the rapid sequestration of chaperones by misfolded proteins accumulating during stress signals the activation of an eIF2 kinase (see **Figure 2**). Although the subcellular location and degree of protein misfolding determine whether ER resident (GRP) or cytoplasmic (HSP) chaperones are eventually induced, significant misfolding in either subcellular site usually signals increased eIF2 phosphorylation and translational suppression within minutes.

Interruption of ER Protein Processing

ER stress results from an accumulation of underprocessed proteins within the organelle that are incompetent of ER to Golgi transport. Accumulation is associated with a variety of treatments that inhibit various processing steps and with genetic variations that result in the expression of unprocessible proteins. Depletion of ER-sequestered Ca^{2+} in response to extracellular Ca^{2+} chelators, Ca^{2+} ionophores, or inhibitors of ER Ca^{2+} uptake interrupts the processing of sugar residues on newly synthesized glycoproteins. Ca^{2+} depletion may also impair the folding of various proteins and the assembly of multimeric proteins in the ER. Mild reducing agents impede the formation of critical disulfide bonds within proteins by shifting the lumenal redox potential toward a less oxidizing state such that folding intermediates accumulate. Concentrations of Ca^{2+} depleting agents or reducing agents that are sufficient to inhibit processing do not lower cellular ATP; removal of the agent by washing results in almost immediate restoration of processing and amino acid incorporation. ER stress is also precipitated by the expression of genetically altered unprocessible proteins and by the unbalanced (nonstoichiometric) production of the subunits of

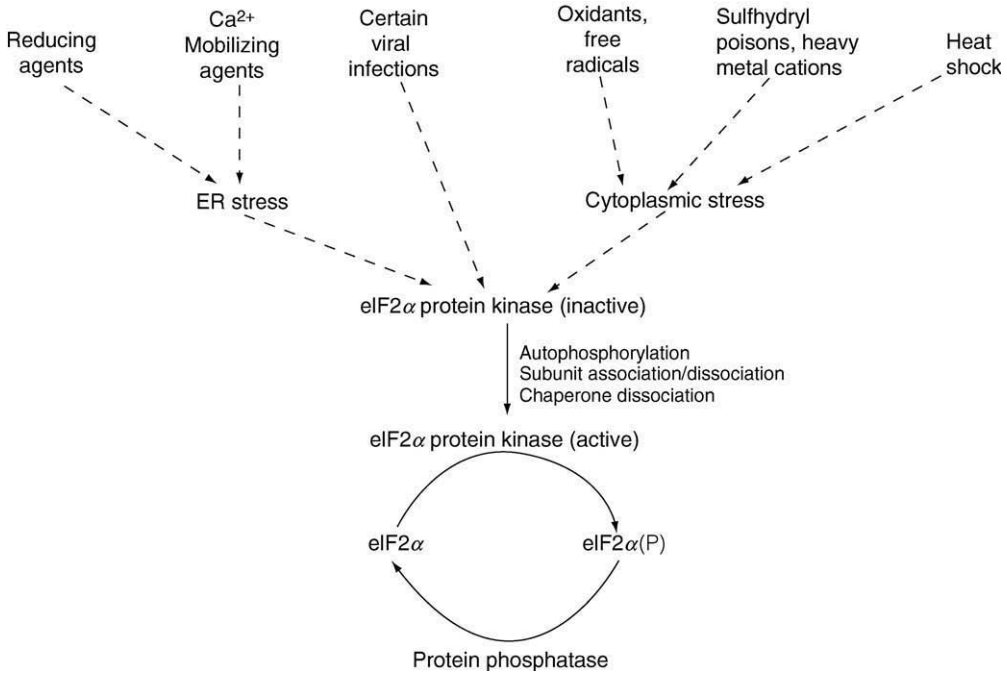


Figure 2 Activation of eIF2α protein kinase by stress.

multimeric proteins. Tunicamycin, which inhibits asparagine-linked protein glycosylation, and brefeldin A, which produces retrograde coalescence of the Golgi with the ER, result in slowly developing ER stress. These various disturbances of ER function inhibit translational initiation coincident with the phosphorylation of eIF2. Phosphorylation has been ascribed to the activation of PKR and PERK.

A stress response characterized by the increased synthesis of new GRP78 and GRP94 molecules is routinely observed on continued (hours) ER stress. The sequestration of ER chaperones by the underprocessed proteins retained within the organelle is thought to initiate the signaling for both translational suppression through eIF2 kinase activation and the induction of new chapterones, albeit by different pathways. The induction of grp78 mRNA is dependent on the activation of an ER membrane-spanning protein kinase termed Irel. The ER lumenal domains of PERK and Irel possess high homology. It is likely that both Irel and PERK are inhibited by interactions with GRP78.

Protein Misfolding in the Cytoplasm

Mammalian cells also undergo an acute suppression of mRNA translation during moderate elevations of their optimal ambient temperature or challenge with certain chemical agents. These treatments, which have been termed proteotoxic, promote mis-folding or damage of cytoplasmic proteins. Proteotoxic agents include sulfhydryl poisons such as sodium arsenite and iodoacetamide, heavy metal cations, and chemical oxidants and free radical generators. The inhibition of mRNA translation in response to most proteotoxic conditions is accompanied by the binding of HSP chaperones to misfolded or damaged proteins, elF2 phosphorylation, and slowing of translational initiation. Following the initial insult, cells typically express abundant quantities of HSPs of 28–30, 60, 70–72, 90, and 110 kDa mass over the next several hours. These inductions collectively are frequently termed the heat shock response and are accompanied by dephosphorylation of elF2 and resumption of translation.

The phosphorylation of elF2 in response to various types of cytoplasmic stress in increased to variable degrees ranging from pronounced phosphorylations with sodium arsenite, low to high phosphorylations with heat stress, and marginal phosphorylations with iodoacetamide. These variations relate to the difficulties in achieving good experimental reproducibility, utilizing conditions that damage cells in a relatively indiscriminate fashion. More stringent conditions that generate relatively complete translational inhibition tend to reduce cell viability, whereas a less rigorous challenge produces only partial suppression. HSP-inducing chemicals observed to promote a rapid phosphorylation of elF2 include sodium arsenite, Cd^{2+}, Hg^{2+}, *t*-butylhydroperoxide, menadione, and diamide. Arsenite has proven particularly valuable for investigating the regulation of protein synthesis

during cytoplasmic stress as it does not produce ER stress or lower ATP. Reproducible, strong stress responses are produced by the drug as are good recoveries of translational activity during periods of increased HSP expression. While arsenite inactivates many proteins with sulfhydryl groups, inactivation is repaired more readily than that produced by high temperatures, heavy metals, or strong oxidants. Arsenite suppresses translational initiation to degrees that correlate closely with the activation of PKR and the phosphorylation of elF2. Peptide chain elongation is not affected, energy stores are not depleted, and viability is maintained.

The cytoplasmic chaperone HSC70 interacts with nascent polypeptides in nonstressed cells, most probably to prevent premature folding. In stressed preparations, nascent polypeptides unable to fold properly and mature proteins that experience damaged folding become stably bound to HSC70. Sequestration of this chaperone during cytoplasmic stress, therefore, has been proposed to signal the heat shock response. Inactivation of PKR has been reported to involve interaction of HSP70 with P58, a cellular inhibitor of the kinase. Various HSPs also function prominently in steroid hormone delivery to the nucleus, the regulation of chaperone-dependent protein kinase activities, clathrin uncoating of coated vesicles, and the transport of protein into the mitochondria.

Nutritional Stress and Stress-Related Apoptosis

Amino acid starvation is associated with slowed rates of translational initiation and, at least in yeast, increased elF2 phosphorylation. The mechanisms signaling phosphorylation under these conditions are unknown, although the depletion of active chaperones may be involved. For example, GRP78 of cultured cells is known to be inactivated by ADP-ribosylation after nutrients in growth media are expended and protein synthetic rates are slowed, elF2 phosphorylation and translational suppression, attributed to the activation of PKR, have been observed in cells undergoing programmed cell death in response to inflammatory cytokines, growth factor deprivation, and ultraviolet light. The extent to which proteins misfold and stress proteins are induced under any of the just-described conditions is unclear. However, glucose deprivation promotes expression of ER chaperones in cultured cells, presumably as a consequence of decreased protein glycosylation.

Stress and the Activation of HRI

HRI of reticulocyte lysates is activated in response to heat shock, oxidative stress, and addition of denatured proteins. In lysates the kinase is known to associate with multiple stress chaperones, including HSP90 and HSC70. HSP90 protects HRI during its synthesis and stabilizes the mature kinase such that it can be autophosphorylated and converted to an activatable form without undergoing stress-dependent denaturation. Because HSP90 dissociates on conversion of the kinase to the activatable form, the chaperone does not function in the process by which stress promotes activation. HSC70 also interacts with nascent and mature competent HRI. Unlike HSP90, however, HSC70 remains associated with the fully activatable kinase. Dissociation of this chaperone fosters activation whereas association suppresses activation. In the intact stressed erythroid cell, therefore, sequestration of HSC70 by misfolded proteins may serve to signal activation of HRI, resulting in elF2 phosphorylation and slowing of protein synthesis. Because development of erythroid stem cells into reticulocytes involves the ejection of the nucleus in conjunction with loss of functional ER, increased expression of ER or cytoplasmic chaperones is not observed in the more differentiated red cell during stress.

The central role of elF2 kinase in the translational suppression that transpires during viral infection, ER stress, or cytoplasmic stress is modeled in **Figure 2.** Using this model, misfolded proteins accumulating during stress compete with the kinase for chaperone binding. Dissociation of chaperones, in combination with subunit rearrangements, promotes autophosphorylation and activation of the kinase. elF2α is phosphorylated and initiation is slowed. Reversibility of translational inhibition requires reassociation of chaperones to repress kinase activity and a protein phosphatase to dephosphorylate elF2α.

Comparable Roles of the Endoplasmic Reticulum and Cytoplasmic Stress Responses in Recovery of Translational Activity

Separate, but closely related, stress response systems exist for the ER, regarding the induction of the GRPs, and for the cytoplasm, pertaining to the induction of the HSPs. As discussed earlier, mRNAs for GRPs and HSPs possess unusual structural features at their 5′ ends thought to favor their selective translation during stress. GRP78 and GRP94, the most prominently induced ER stress chaperones, share significant sequence homology with HSP70 and HSP90, respectively. GRP78 is hypothesized to function in early ER protein folding and assembly, in the translocation of proteins from the cytosol to the ER for processing,

and in the retention and degradation of improperly folded proteins. Comparable chaperone activity is exhibited by HSP70 and a closely related form of the protein expressed in nonstressed cells, HSC70. Both prevent incorrect folding of polypeptides during synthesis, permitting delivery to organelles in an unfolded state for translocation. These chaperones may also solubilize or refold denatured or aberrant proteins and/or deliver them to a degradative system. Induction of GRPs or HSPs appears necessary for survival during persistent ER or cytoplasmic stress, respectively.

The induction of either GRPs or HSPs in stressed cells is accompanied by a reduced phosphorylation of elF2 and a partial resumption of mRNA translation. Induction of either class of stress proteins is sufficient for the development of translational tolerance to subsequent rechallenge of the cells with either ER or cytoplasmic stressors. Cross-tolerance is observed in terms of continued amino acid incorporation into proteins, maintenance of polyribosomal contents, and the lack of increased elF2 phosphorylation. Translational recovery from inhibition by Ca^{2+}-releasing drugs is partly overturned by antisense oligonucleotides directed against *grp78* mRNA, whereas overexpression of GRP78 is sufficient for tolerance to translational inhibition and elF2 phosphorylation in response to ER stress. Presumably various GRP and HSP chaperones inhibit elF2 kinase(s) through complexing with a critical component(s) of the enzymes or with another protein(s), which affects activity. If so, the sequence homology of GRP78 and HSP70/HSC70 may be important to the putative dual input.

By regulating elF2 kinase activity, the two stress chaperone systems permit rates of protein folding or processing to be coordinated with rates of protein synthesis through changes in the rate of recycling of elF2 and of ribosomal loading onto mRNA. The interactions of GRP78 in regulating translational initiation are modeled in **Figure 3**.

With the exception of their mutual abilities to influence elF2 phosphorylation, the two stress systems appear to function independently. For example, sodium arsenite does not induce the GRPs nor does it affect Ca^{21} retention by the ER. Following induction of GRP78 with Ca^{2-}-mobilizing agents, cells remain responsive to HSP induction by subsequent arsenite treatment. Similarly, cells expressing the cytoplasmic stress response remain susceptible to induction of the GRPs. It should be recognized that

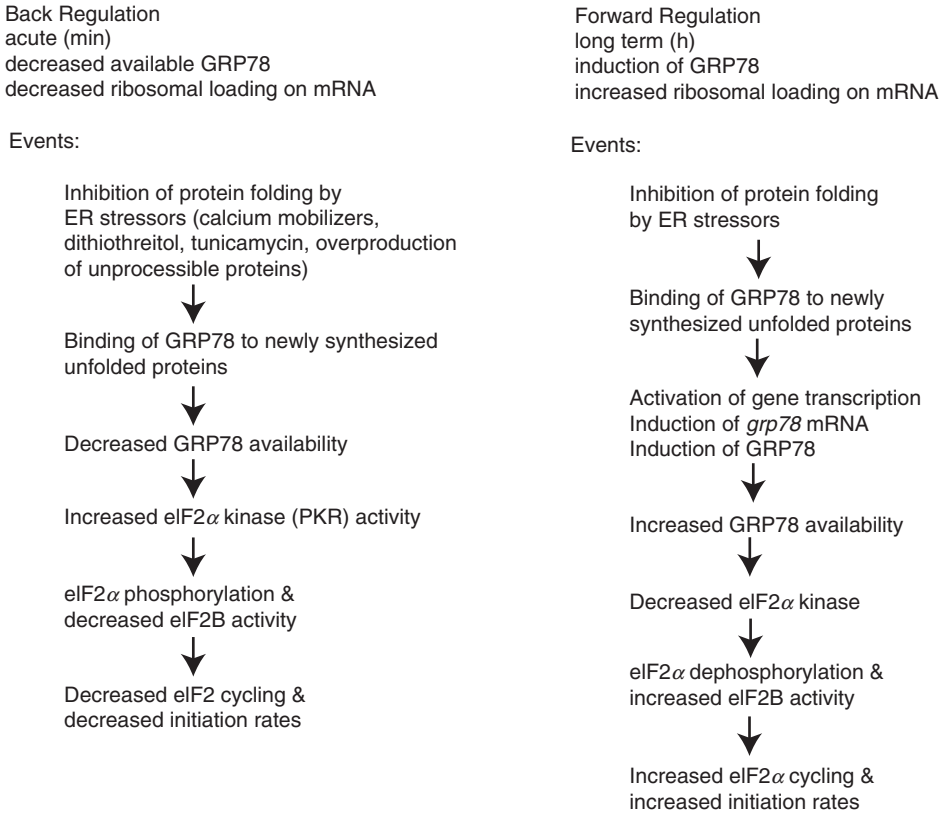

Back Regulation
acute (min)
decreased available GRP78
decreased ribosomal loading on mRNA
Events:
Inhibition of protein folding by ER stressors (calcium mobilizers, dithiothreitol, tunicamycin, overproduction of unprocessible proteins)
Binding of GRP78 to newly synthesized unfolded proteins
Decreased GRP78 availability
Increased elF2α kinase (PKR) activity
elF2α phosphorylation & decreased elF2B activity
Decreased elF2 cycling & decreased initiation rates
Forward Regulation
long term (h)
induction of GRP78
increased ribosomal loading on mRNA
Events:
Inhibition of protein folding by ER stressors
Binding of GRP78 to newly synthesized unfolded proteins
Activation of gene transcription
Induction of grp78 mRNA
Induction of GRP78
Increased GRP78 availability
Decreased elF2α kinase
elF2α dephosphorylation & increased elF2B activity
Increased elF2α cycling & increased initiation rates

Figure 3 Coordination of rates of protein processing and protein synthesis.

the phosphorylation of elF2 and the inhibition of translation are not mandatory for the induction of stress proteins.

Low concentrations of ER stressors clearly induce GRP78 in the absence of elF2 phosphorylation or inhibition of translation. Greater degrees of cytoplasmic stress also appear to be required for the signaling of translational suppression than for induction of the HSPs. These findings are consistent with the existence of pathways to slow the synthesis of protein precursors only when processing or folding is seriously impaired.

eIF2 Kinase as a Focal Point for Regulating Protein Synthesis during Stress

The body of available information with nonerythroid mammalian cells suggests that PERK and PKR function together in governing rates of translational initiation in response to various stimuli including, but not limited to, ER and cytoplasmic stresses. HRl, a heme-regulated protein kinase, is unlikely to mediate translational suppression in stressed nonerythroid cells. Because the induction of either HSPs or GRP78 provides cross-tolerance to either ER or cytoplasmic stressors, it seems clear that PERK and PKR interact in some manner, perhaps through cross-dimerization of their respective kinase domains during activation. It is, of course, quite possible that additional homologous forms of elF2 kinase remain to be discovered. Putative additional stimuli for elF2 kinase activation could include hormonal or nutritional alterations or treatments that damage cell membranes.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Chaperonopathies; Viral Virulence and Stress.

Further Reading

Brostrom, C. O. and Brostrom, M. A. (1998). Regulation of translational initiation during cellular responses to stress. *Progress in Nucleic Acid Research and Molecular Biology* **58**, 79–125.

Brostrom, M. A. and Brostrom, C. O. (2003). Calcium dynamics and endoplasmic reticular function in the regulation of protein synthesis: implications for cell growth and adaptability. *Cell Calcium* **34**, 345–363.

Clemens, M. J. (1996). Protein kinases that phosphorylate elF2 and elF2B, and their role in eukaryotic cell translational control. In: Hershey, J. W. B., Mathews, M. B. & Sonnenberg, N. (eds.) *Translational control*, pp. 139–172. Plainview, NY: Cold Spring Harbor Laboratory Press.

Clemens, M. J. and Elia, A. (1997). The double-stranded RNA-dependent protein kinase PKR: Structure and function. *Journal of Interferon and Cytokine Research* **17**, 503–524.

De Gracia, D. J. and Montie, H. L. (2004). Cerebral ischemia and the unfolded protein response. *Journal of Neurochemistry* **91**, 1–8.

de Haro, C., Mendez, R. and Santoyo, J. (1996). The elF-2α kinases and the control of protein synthesis. *FASEB Journal* **10**, 1378–1387.

Duncan, R. F. (1996). Translational control during heat shock. In: Hershey, J. W. B., Mathews, M. B. & Sonnenberg, N. (eds.) *Translational control*, pp. 271–293. Plainview, NY: Cold Spring Harbor Laboratory Press.

Gimenez-Barcons, M., Wang, C., Chen, M., et al. (2005). The oncogenic potential of hepatitis C virus NSSA sequence variants is associated with PKR regulation. *Journal of Interferon and Cytokine Research* **25**(3), 152–164.

Harding, H. P., Zhang, Y. and Ron, D. (1999). Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature* **397**, 271–274.

Hershey, J. W. B. (1991). Translational control in mammalian cells. *Annual Review of Biochemistry* **60**, 717–755.

Melville, M. W., Tan, S., Wambach, M., et al. (1999). The cellular inhibitor of the PKR protein kinase, $P58^{IPK}$, is an influenza virus-activated co-chaperone that modulates heat shock protein 70 activity. *Journal of Biological Chemistry* **274**, 3797–3803.

Pain, V. M. (1994). Translational control during amino acid starvation. *Biochimie* **76**, 718–728.

Pain, V. M. (1996). Initiation of protein synthesis in eukaryotic cells. *European Journal of Biochemisty* **236**, 747–771.

Panniers, R. (1994). Translational control during heat shock. *Biochimie* **76**, 737–747.

Pestova, T. V., Borukhov, S. I. and Hellen, U. T. (1998). Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature* **394**, 854–859.

Proud, C. G. (1992). Protein phosphorylation in translational control. *Current Topics in Cell Regulation* **32**, 243–369.

Welch, W. J. (1993). Heat shock proteins functioning as molecular chaperones: Their roles in normal and stressed cells. *Philosophical Transactions of the Royal Society of London B* **339**, 327–333.

Williams, B. R. (1997). Role of the double-stranded RNA-activated protein kinase (PKR) in cell regulation. *Biochemistry Society Transactions* **25**, 509–513.

Wiseman, R. L. and Balch, W. E. (2005). A new pharmacology—drugging stressed folding pathways. *Trends in Molecular Medicine* **11**, 347–350.

Protein Translocation *See:* Protein Synthesis; Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Response, Overview; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Proteosome

M Maldonado
University of Pennsylvania, Philadelphia, PA, USA
J Wang
Wyeth Research, Collegeville, PA, USA

© 2007 Elsevier Inc. All rights reserved.

The Ubiquitin-Proteosome System
The Ubiquitin-Proteosome System in the Pathogenesis of Disease
Conclusion

Glossary

26S Proteosome	Large protease complex composed of a 20S proteolytic-containing complex capped by two 19S regulatory particles responsible for degradation of polyubiquitinated proteins.
E3 ubiquitin ligase	Diverse group of enzymes with distinct motifs for recognizing specific substrates that carry out the final transfer of ubiquitin to the bound substrate.
N-end rule	Relates the N-terminal residue to the half-life of an intracellular protein.
Ubiquitin-proteosome system	Major proteolytic pathway for intracellular protein degradation involved in many key cellular processes.
Ubiquitination	Enzymatic process that sequentially covalently links ubiquitin to targeted protein substrates to form a polyubiquitin chain recognized by the 26S proteosome.
Unfolded protein response	Highly specific signaling pathway in the endoplasmic reticulum (ER) that recognizes the accumulation of un- or misfolded proteins caused by disruption of ER homeostasis.

Understanding cellular homeostasis requires knowledge of both synthetic and degradation pathways. Protein degradation, in particular, plays a key regulatory process for many cellular processes. There are two systems involved in intracellular protein degradation. The lysosomal pathway was first identified, but in the late 1950s, Goldberg and colleagues demonstrated a new type of proteolytic system in lysosome-free reticulocyte cell extracts. This led to the discovery of ubiquitin and the proteosome and the elucidation of the ubiquitin-proteosome system (UPS). The UPS degrades the vast majority (up to 80–90%) of long- and short-lived normal and abnormal intracellular proteins, hence is considered to be the major pathway for intracellular protein degradation. A growing body of evidence shows that this is a highly regulated yet complex system that is central to normal homeostasis, including cell cycle regulation, DNA repair, regulation of immune and inflammatory responses, and cellular response to stress. Proteins targeted by the UPS include mutated and misfolded proteins, cyclins, and tumor suppressors. Derangements of the UPS can lead to many disorders, including malignancies, neurodegenerative diseases, and possible systemic autoimmunity. A better understanding of the UPS process and identification of the components involved in the degradation of key regulatory proteins have led to the development of mechanism-based therapeutics in various diseases.

The Ubiquitin-Proteosome System

Ubiquitin and Ubiquitination

The control of protein turnover must logically, but not necessarily intuitively, be a counterpart to protein synthesis. This dynamic balance, revealed by variability in the degradation rates and half-life of different proteins, relies on the UPS. The proteolysis is energy dependent and specific. Proteins are selectively targeted for degradation by covalent linkage to ubiquitin, a 76-residue protein that is highly evolutionarily conserved in all eukaryotes. The name ubiquitin is a misnomer, having arisen from the original belief that it was found in all cellular organisms. The linkage of ubiquitin to target protein is followed by sequential addition of more ubiquitin moieties to each other to form a polyubiquitin chain that functions as the

recognition signal for a downstream proteosome in the UPS. Poly- or monoubiquitination can be involved in a variety of nonproteolytic functions, such as activating enzymes, modification of histones that regulate transcription, modulation of membrane dynamics, or routing of the tagged proteins to their subcellular compartments. Ubiquitin-binding proteins that have ubiquitin-binding domains recognize some ubiquitinated proteins. Modification of target protein by ubiquitin or a ubiquitin-like protein remodels the surface of target proteins, affecting their stability, interactions with other proteins, activity, and subcellular localization. At least 10 different modifiers have been described in mammalian cells, and conjugation of each modifier to its target has different biological effect.

Ubiquitination Reaction

Ubiquitination of cellular proteins is a highly complex, temporally controlled, and tightly regulated energy-dependent process. The conjugation of ubiquitin is carried out in a cascade reaction involving three enzymes designated E1, E2, and E3 (Figure 1). Ubiquitination begins with the attachment of ubiquitin by E1, also known as the ubiquitin-activating enzyme. Ubiquitin is then transferred to E2, or the ubiquitin-conjugating enzyme. The final step is dependent on one of many E3s (several hundred forms have been identified in the human genome) that are responsible for catalyzing the addition of ubiquitin to the protein substrate. Chain elongation, or polyubiquitination, can be carried out by the same E3 or by a different ligase, which can be referred to as E4.

Substrate Targeting

The UPS is highly specific and selective. Much of this lies in the diversity of different ubiquitin-protein ligase E3s that can recognize a specific substrate. Many substrates are recognized only after modifications such as misfolding, denaturing, hydroxylation, or phosphorylation. Molecular chaperones such as heat shock protein 70 or chaperon-like molecules such as valosin-containing protein can also act as recognition elements. E3 ligases recognize specific targeting motifs, which can vary from a single amino acid residue to a domain not usually exposed. One of the most important recognition patterns is a destabilizing N-terminal amino acid. A sequence such as arginine and lysine can be a unique feature of a protein, which relates the half-life of an intracellular protein to its N-terminal residue, which has become known as the N-end rule. Phosphorylation of either the substrate or the ligase can also control targeting. Ubiquitin-like proteins can also play a regulatory role by protecting proteins from ubiquitination.

In addition to its substrates, the activity of the ubiquitin system itself can be modulated by a number

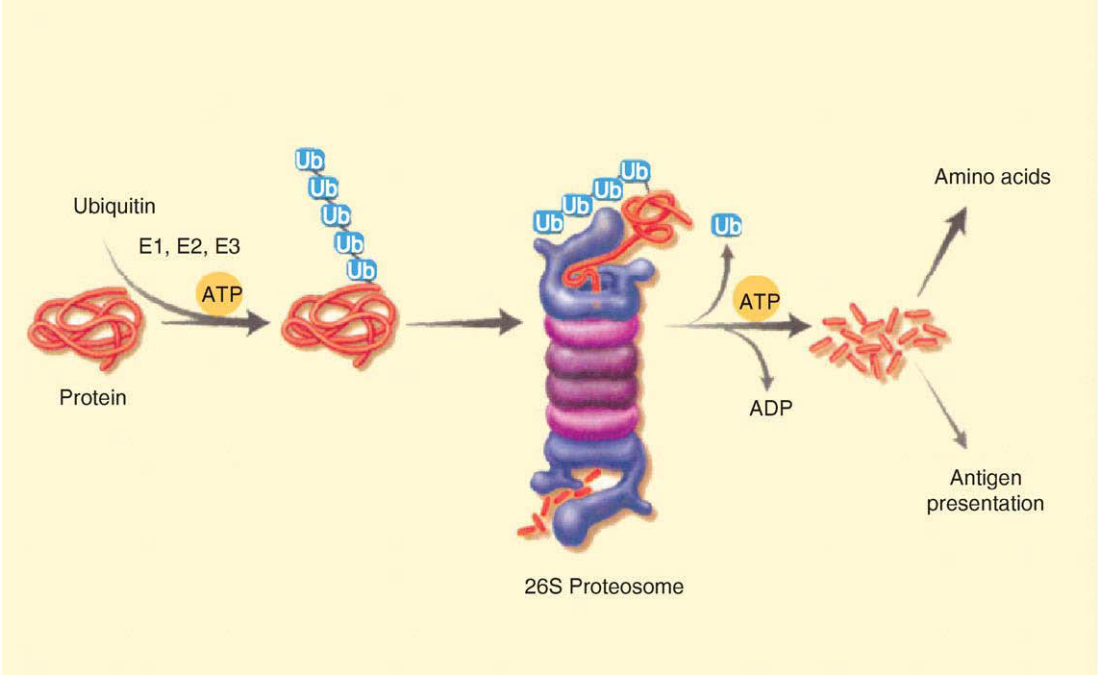


Figure 1 Schematic of the ubiquitin-proteosome system. Ubiquitin is added sequentially to targeted substrates in an energy-dependent enzymatic cascade requiring three enzymes, E1, E2 and E3. Polyubiquitinated proteins are recognized by the 26S proteosome. Cleavage produces small peptides and free ubiquitin. Adapted from Kisselev and Goldberg (2001) with permission from Elsevier.

of factors, including thyroid hormones, glucocorticoid steroids, cytokines, and proteins expressed in malignant cells such as proteolysis-inducing factor (PIF). Interestingly, some factors such as interferon gamma (IFN-γ) not only regulate the modification of substrate, including the IκB family of proteins, but also modify the components of the enzymatic machinery of the ubiquitin system and the proteosome complex. Similarly, protein kinase C and the tyrosine kinase pathway are involved in both the modification of substrates of UPS and phosphorylation of E1 and E2, thereby exerting their activities with high efficiency.

The Proteosome

The ubiquitinated substrates targeted for proteolysis are recognized by a 26S complex cellular structure, the proteosome (Figure 2). This high-molecular-mass protease is also ATP dependent. Degradation of proteins by proteosomes removes denatured, damaged, or improperly translated proteins and regulates the level of proteins in the cells. The 26S proteosome is composed of two subcomplexes. One or two 19S regulatory particles (RPs) cap either or both ends of the barrel-shaped 20S complex, which carries the proteolytic activity, forming 26S and 30S proteosomes, respectively. The 19S RP is responsible for recognizing the ubiquitinated proteins as proteosome substrates, the opening of the orifice in the 20S complex that allows entry onto the proteolytic chamber and likely the unfolding of the substrate necessary for entry. Within the chamber, proteins are broken down into short peptides. Without the 19S RP, the 20S complex is extremely inefficient in degrading folded protein substrates. The 20S complex is composed of four stacked rings, two identical outer α rings and two inner β rings. Each ring has seven distinct subunits. The β rings contains multiple catalytic sites including chymotrypsin-like sites that are the target of almost all the synthetic and natural inhibitors of the proteosome. The outer two α rings have no known function. Other subcomplexes can cap the 20S complex, creating a polar arrangement that can affect the product generated.

Degradation Products

After the protein has been degraded, the 19S RP removes the polyubiquitin tag from the substrate protein. Degradation in the proteosome generates short peptides that can have different fates. Some are further degraded by cytosolic peptidases. Some can be transported into the endoplasmic reticulum (ER) for binding to MHC class I molecules and cell surface expression. This latter subset of antigenic peptides may arise from proteosome activator PA28 subcomplex capped polar proteosomes that generate peptides trimmed to properly fit the MHC class I binding grooves. In this regard, hybrid proteosomes (19S

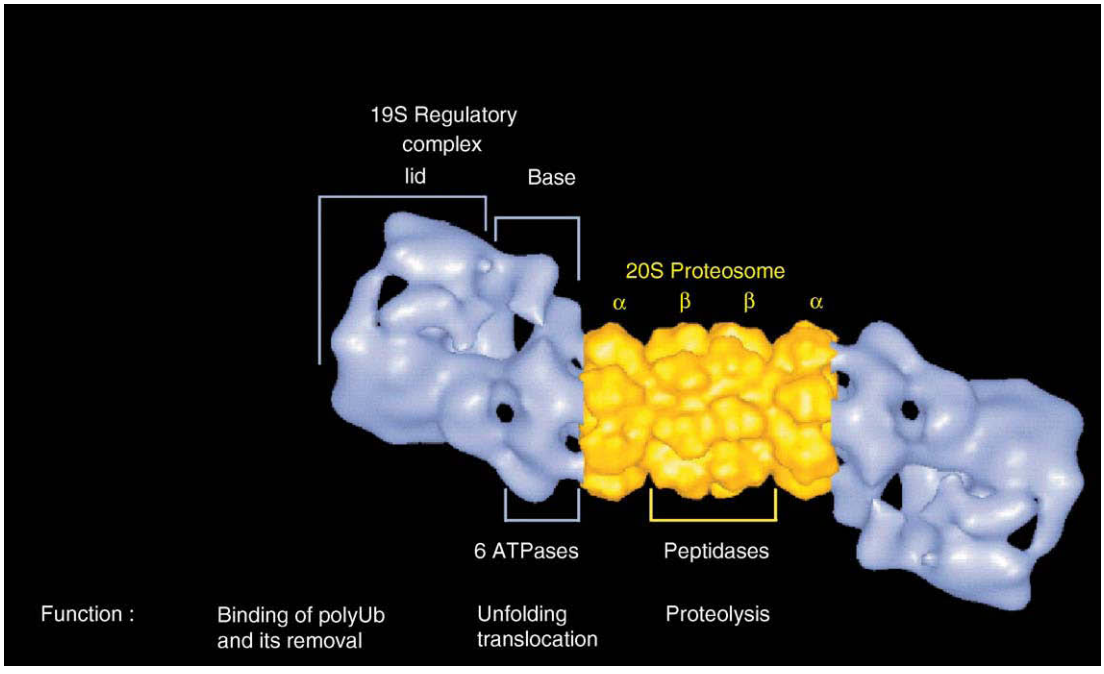


Figure 2 Electron tomographic image and structure of the 26S proteosome microscope. The 26S proteosome is composed of the barrel-shaped 20S complex capped by two 19S subcomplexes. The 19S regulatory particle recognizes the polyubiquitin tag on targeted substrates and unfolds the substrate to enter the proteolytic chamber. The 20S core particle contains the catalytic sites responsible for the proteolysis. Adapted from Kisselev and Goldberg (2001) with permission from Elsevier.

RP–20S proteosome–PA28) may have unique value during immune responses. Peptides are not always the final product of the UPS, as longer, biologically active proteins can be generated from precursors. Free, reusable ubiquitin is also generated. Deubiquitylating enzymes catalyze this reaction.

The Ubiquitin-Proteosome System in the Pathogenesis of Disease

The Role of the UPS in Stress and Disease

Regulation of protein turnover has an important role in various cellular functions, and alterations can lead to many diseases. Viruses have evolved proteins that can exploit this system to their advantage by either enhancing or avoiding ubiquitination. The unfolded protein response (UPR) is a highly specific signaling pathway in the ER that recognizes the accumulation of un- or misfolded proteins caused by disruption of ER homeostasis. Activating the UPR has multiple effects, including enhancing degradation of the luminal content of the ER via the UPS, and has been implicated in numerous disease processes.

Malignancies

Protein degradation via the ubiquitin-proteosome system is an important regulator of cell cycle progression and survival through modulating the function of nuclear factor-κB (NF-κB) and pro-apoptotic transcription factor p53. Thus, inhibition of this pathway can be directly cytotoxic and can sensitize several tumor cell types to cytotoxic chemotherapy and radiation. Most of the experience with proteosome inhibition has been gained through work done in this disease area. Preclinical studies have demonstrated that the proteosome inhibitor bortezomib decreases proliferation, induces apoptosis, enhances the activity of chemotherapy and radiation, and reverses chemoresistance in a variety of hematologic and solid malignancy models *in vitro* and *in vivo*. Bortezomib was the first proteosome inhibitor to enter clinical trials. In two phase II trials, bortezomib alone or in combination with dexamethasone elicited durable responses with meaningful survival benefits in patients with recurrent and/or refractory multiple myeloma. In a phase III trial, bortezomib produced significant survival benefits and improved response rates over high-dose dexamethasone at first recurrence and beyond in patients with multiple myeloma. Current studies indicate that bortezomib may serve as induction therapy before stem cell transplantation. In addition, proteosome inhibition with bortezomib also has shown activity and manageable toxicity in mantle cell and other lymphoma.

Neurodegenerative Diseases

Dysregulation of the UPS has been implicated in the pathogenesis of both inherited and acquired neurodegenerative diseases. Accumulation of ubiquitin conjugates and/or inclusion bodies associated with ubiquitin, proteosomes, and certain disease-specific proteins has been demonstrated in a wide variety of neurodegenerative diseases. The list includes neurofibrillary tangles of Alzheimer's disease, brain stem Lewy bodies, the pathological hallmark in Parkinson's disease, Bunina bodies in amyotrophic lateral sclerosis, and nuclear inclusions in CAG repeat expansion disorders such as Huntington's disease, spinocerebellar ataxias, and spinal and bulbar muscular atrophy (Kennedy's disease). This phenomenon is thought to be due to failed attempts of the UPS to process damaged and abnormal proteins, or movement of the proteosome to specific subcellular sites where degradation of abnormal proteins may occur.

There are at least three key players that link aberration in the UPS to the pathogenesis of neurodegenerative diseases. The first is Parkin, a 52-kDa RING finger-containing E3 ligase that ubiquitinates and promotes degradation of several proteins related to the pathogenesis of Parkinson's disease. Mutations in the gene have been found in ~50% of patients with autosomal recessive juvenile Parkinsonism, one of the most common familiar form of Parkinson's disease. Second is α-synuclein, a 140-amino-acid residue protein shown to be a major component of Lewy bodies and Lewy neuritis in sporadic Parkinson's disease, dementia with Lewy bodies, and the Lewy body variant of Alzheimer's disease. Third is ubiquitin C-terminal hydrolase (UCH-L1). Mutations of UCH-L1 were discovered in a German family with Parkinson's disease.

The most compelling evidence for the involvement of the UPS in the pathogenesis of Alzheimer's disease derived from the discovery that a frameshift mutation in the ubiquitin transcript resulted in a mutant form of ubiquitin called UBB^{+1}. This mutant has been observed in brains of Alzheimer's disease patients. Expression of UBB^{+1} in the brain could inhibit the degradation of a polyubiquitinated substrate by the 26S proteosome, leading to accumulation of toxic proteins with neuropathological consequences.

Inflammation and Autoimmune Diseases

Many autoimmune diseases are characterized by immunologically mediated organ damage. Some of these autoimmune diseases may have recognizable environmental triggers, such as infections, but it is not known how or why inflammation is initiated in most and sustained in all. Key to these diseases is the same cells and molecules, such as cytokines, that

are also central to normal immunity. The UPS is now known to regulate many of these, including MHC-mediated antigen presentation, cytokine regulation, and cell cycle, and must therefore also have a role in their pathogenesis and be a logical therapeutic target.

One of the prominent examples for the involvement of the UPS in inflammation is MHC class I antigen processing by antigen-presenting cells (APCs). In addition, the UPS plays a significant role in the regulation of both T cell receptor (TCRs) and costimulatory CD28 signaling through the action of ubiquitin ligases of the Cbl family. CD28 costimulation results in the ubiquitination and degradation of Cbl-b, which eliminates the negative regulators and allows the expression of proinflammatory cytokines and their receptors. However, the most important link between the UPS and inflammation is related to NF-κB. NF-κB is a master regulator of many inflammatory cytokine genes, and its activation is mediated through the UPS. NF-κB is actively inhibited when bound to IκB. NF-κB activation follows the degradation of IκB, which is dependent on ubiquitination of IκB followed by proteosomal degradation. Hence, alterations in the UPS have profound effects on immune responses, including the regulation of an array of inflammatory cytokines.

INF-γ is a key pleiotropic regulatory cytokine. It controls an inducible proteolytic cascade that provides peptides for MHC class I presentation. Seronegative spondyloarthropathies (SpA) are a group of diseases characterized by, but not limited to, axial joint inflammatory. Ankylosing spondylitis (AS) is the prototypical SpA. Most patients with AS carry the MHC class I HLA-B27 gene, so its role in disease pathogenesis is an area of intense research. Much interest has been focused on determining the origin and nature of the peptides being presented by HLA-B27 and the cell surface expression of misfolded HLA-B27, key areas in which the UPS is known to play a role.

The UPS is involved in the regulation or induction of apoptosis. Apoptosis has been implicated in both experimental models and patients with systemic lupus erythematosus. Altered clearance of autoantigens is thought to allow for targeting by the immune system and the development of autoimmunity.

The UPS may be involved in the pathogenesis of multiple autoimmune diseases through many different mechanisms. It may therefore be a therapeutic target worth considering. Conversely, targeting the UPS may produce unwanted immunological side effects.

Cardiovascular Diseases

Inflammation, cell proliferation, and apoptosis constitute major feature of atherosclerosis. The UPS is involved in the development of atherosclerosis through its regulation of these biological pathways. In addition, the UPS is associated with the pathogenesis of acute coronary syndromes through diverse cellular and intracellular targets and modes of actions. Recent experimental evidence indicates that the UPS is functionally active in early atherogenesis despite increased oxidative stress mimicking the initiation stage of atherosclerosis. The pathophysiological role of the UPS in the proliferative aspects of cardiovascular disease was demonstrated by the intense ubiquitin nuclear staining in intimal thickening lesions. Treatment of cultured vascular smooth muscle cells with proteosome inhibitor has led to a dose-dependent inhibition of proliferation and induction of apoptosis. In rodent models, single-dose local administration of the proteosome inhibitor MG-132 for 5 min after endothelial denudation of carotid arteries reduces neointima formation by 75%. Further, proteosome inhibition led to significant reduction in vascular atrophy in DOCA-salt hypertensive rats.

T cells have been identified as a characteristic inflammatory element of atherosclerosis, which accumulate at the rupture-prone sites of the plaque shoulder regions. The UPS is involved in the pathogenesis of atherosclerosis by regulating the activation of T cells through T cell receptors and costimulatory molecule CD28. The UPS may be involved in foam cell formation as well. In an autopsy-based study, increased expression of ubiquitin/ubiquitin conjugates was colocalized with macrophages and apoptotic cells in the area of the lipid core. It has been shown that the expression of p53 in human atherosclerotic plaques was enhanced, and overexpression of p53 resulted in a significant increase in apoptosis and decrease in extracellular matrix in the cap region. Accumulating evidence indicates a strong association of pro-apoptotic protein p53, ubiquitin, and apoptotic neointimal cells in the cap region of unstable plaques from patients with fatal myocardial infarction.

Conclusion

The UPS is a complex system that controls many important aspects of cell function. The covalent linkage of ubiquitin to protein substrates can selectively and specifically alter their fate through proteolytic and nonproteolytic pathways. Loss of normal homeostasis through many mechanisms, including stress, can lead to aberrant cellular function and disease.

See Also the Following Articles

Autoimmunity; Chaperone Proteins and Chaperonopathies; Chaperonopathies; Heat Shock Response, Overview;

Homeostasis; Inflammation; Neurodegenerative Disorders; Parkinson's Disease; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles; Protein Synthesis.

Further Reading

Adams, J. (2003). The proteasome: structure, function, and role in the cell. *Cancer Treatment Reviews* **29** (Supplement I), 3–9.

Ciechanover, A. (2003). The ubiquitin proteolytic system and pathogenesis of human diseases: a novel platform for mechanism-based drug targeting. *Biochemical Society Transactions* **31**, 474–481.

Ciechanover, A. (2005). Intracellular protein degradation: from a vague idea thru the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Cell Death and Differentiation* **12**, 1178–1190.

Ciechanover, A. and Brundin, P. (2003). The ubiquitin proteosome system in neurodegenerative diseases: sometimes the chicken, sometimes the egg. *Neuron* **40**, 427–446.

Colmegna, I., Sainz, B. Jr., Garry, R. F. and Espinoza, L. R. (2005). The proteasome and its implications in rheumatology. *Journal of Rheumatology* **32**, 1192–1198.

Herrmann, J., Ciechanover, A., Lerman, L. O. and Lerman, A. (2004). The ubiquitin-proteosome system in cardiovascular diseases – a hypothesis extended. *Cardiovascular Research* **61**, 11–21.

Kisselev, A. L. and Goldberg, A. L. (2001). Proteasome inhibitors: from research tools to drug candidates. *Chemistry & Biology* **8**, 739–758.

Richardson, P. G., Mitsiades, C., Hideshima, T. and Anderson, K. C. (2006). Bortezomib: proteasome inhibition as an effective anticancer therapy. *Annual Review of Medicine* **57**, 33–47.

Welchman, R. L., Gordon, C. and Mayer, R. J. (2005). Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nature Reviews Molecular Cellular Biology* **6**, 599–609.

Zhang, K. and Kaufman, R. J. (2006). The unfolded protein response: a stress signaling pathway critical for health and disease. *Neurology* **66**(Supplement 1), S102–109.

Psoriasis

A B- Kirschbaum
Technical University of Dresden, Dresden, Germany

© 2007 Elsevier Inc. All rights reserved.

Immunopathology of Psoriasis
Stress – a Triggering Factor in Psoriasis
Psychobiology of Psoriasis
Summary

Glossary

Karatinocytes Major cell type of the epidermis making up to 90% of epidermal cells. Keratinocytes are shed and replaced continuously from the stratum corneum. The time of transit from basal layer to shedding is approximately one month although this can be accelerated in conditions of keratinocyte hyperproliferation, i.e., psoriasis.

Phototherapy Light therapy using specific wavelength of light for a distinct period of time. Generally, lasers, fluorescent light, or full-spectrum light are used.

Superantigens Molecules that stimulate indiscriminately T-cells. The best characterized superantigens are the microbial toxins from *Staphylococcus aureus* and *Streptococcus pyogenes*.

Immunopathology of Psoriasis

Psoriasis (PSO) is a chronic, genetically determined inflammatory skin disease with an estimated prevalence of 2% of the northern population. Apart from allergic contact dermatitis and atopic dermatitis, PSO is the most common chronic skin disease. PSO can start at any age; however, disease onset is most frequently observed in the second decade of life. PSO symptomatology is characterized by erythematous and solid skin plaques with sharply defined borders, covered with fine silvery scales. Although not life threatening, PSO can be a disabling disease with a considerable impact on patients' psychological and social well-being. Psoriatics often feel stigmatized and psychologically stressed by the disfiguring skin disease, which can lead to increased anxiety and depression.

The pathophysiology of PSO is not fully understood. However, accumulating evidence suggests that T-cell-mediated autoimmune processes and the action of pro-inflammatory cytokines cause the hyperproliferation of keratinocytes and assume the psoriatic phenotype. The antigenic peptide responsible for the initial T-cell activation in PSO has not been established yet. However, one current model is that bacterial proteins (streptococcal M proteins) resulting from a preceding infection may act as superantigens. Another important process in PSO pathogenesis

is the trafficking of the activated T cell into the skin. Activated T cells generated from psoriatic skin lesions secrete high concentrations of interleukin (IL-)2, tumor necrosis factor (TNF-)α, and interferon (IFN-)γ, indicating a significant role of T helper 1 (Th_1)-mediated inflammatory processes in the psoriatic skin. The significant role of circulating Th_1-derived cytokines is further supported by the finding that TNF-α and IFN-γ increase the infiltration of T cells and other inflammatory cells such as monocytes into the skin. Locally, IFN-γ stimulates epidermal cell proliferation and keratinocyte hyperplasia, leading to the psoriatic phenotype.

Stress – a Triggering Factor in Psoriasis

Although PSO has been conceptualized as an autoimmune disease primarily mediated by an inappropriate function of T cells, it is broadly accepted that psychosocial stress may affect PSO and may trigger the exacerbation of the disease. For example, 31% of PSO patients reported the onset of PSO in times of increased everyday life stress, and for 71%, PSO symptomatology worsened during stressful life episodes. To further evaluate the role of stress in the onset and exacerbation of PSO, psychocutanous characteristics in PSO sufferers reporting a strong link between psychosocial stress and skin condition (high stress reactors, HSRs) and in PSO patients without any association between stress and PSO symptoms (low stress reactors, LSRs) have been investigated. Clinically, HSRs suffered from a more severe disease, with a distinct distribution of PSO plaques in emotionally charged body regions (i.e., the face, neck, forearms, hands, and genital regions), affecting the patient's physical appearance and sexuality. It has been proposed that exacerbation of the disease may lead to stronger psychological reactions and, thus, may explain the strong psychocutanous relationship in this patient group. Psychologically, the HSRs tended to rely more on the approval of others and scored higher on depressed anger and hostility. Others have reported increased levels of anxiety, aggression, obsessionality, and depression in PSO sufferers, with the last being linked to pruritus and sleep difficulties. It is of note, however, that the personality pattern described in psoriatics can also be identified in patients with atopic dermatitis (AD), a chronic (allergic) inflammatory skin disease characterized by clinical features similar to PSO such as severe pruritus or disfiguring skin symptomatology. These observation suggests that the personality type described in psoriatics may not be a PSO-specific personality type but, rather, represents a personality pattern linked to a chronic inflammatory and disfiguring skin disease. The importance of stress in PSO has been further highlighted by reports suggesting that psychological distress affects treatment outcome in PSO patients. In these studies, it was demonstrated that the level of stress may predict the time it took for photochemotherapy (PUVA: psoralen, P, and long-wave ultraviolet radiation, UVA) to clear PSO symptomatology. Accordingly, stress reduction by stress management, relaxation, or cognitive techniques shortened the time it took to clear PSO symptoms by PUVA and, moreover, improved the clinical severity of PSO.

Psychobiology of Psoriasis

So far we have summarized the immunological and psychological factors crucial in the onset and maintenance of PSO. Based on these findings, it remains to be determined how PSO-relevant psychological and immunological processes are linked and, moreover, by which mechanisms stress may affect skin condition in PSO.

Considerable evidence has emerged showing that the central nervous system (CNS) and the immune system are intimately connected, each regulating the other. One major pathway through which the CNS exerts control over the immune system is via the neuroendocrine system. Neuropeptides released by the brain may act directly on immune cells, or they may affect the release of other hormones from endocrine tissues that can modify immunity. Investigating neuroendocrine-immunity communication pathways in PSO patients and, moreover, studying potential aberrations in the interplay between these systems under stressful conditions may give a hint about how stress can trigger inflammatory processes in PSO and exacerbate PSO symptomatology.

In the last decade, an extensive literature has emerged demonstrating the important immunosuppressive and anti-inflammatory role of the hypothalamus-pituitary-adrenal (HPA) axis. Recently, the role of stress-induced HPA axis activation has been specified – an appropriate reactivity of the HPA axis under stress may be necessary to regulate immune function and to prevent an immune response, for example an inflammatory response, from reaching a level that may be damaging to the host. There are some data suggesting a hyporesponsive or hypoactive HPA axis in PSO patients. For example, attenuated cortisol responses to psychosocial stress or diminished urinary cortisol concentrations in psoriatics have been reported. However, other researchers failed to demonstrate reduced cortisol or adrenocorticotropic hormone (ACTH) responses to stress in PSO sufferers. One factor that may explain the hetereogenous study results may be the acuity or chronicity of the

ongoing inflammatory process in the patient group investigated. There is evidence that normal cortisol levels can be found in chronic nonprogressive PSO characterized by clinical inactivity of chronic plaques, whereas clinically active phases with fresh eruptions were characterized by increased HPA axis activity. These data suggest that it may be useful to consider inflammation in PSO as a dynamic process with different phases of acuity and remission that may be linked to distinct endocrine response patterns.

In general, the increased activity of the sympathoadrenomedullary (SAM) system in PSO subjects has been described. Studies using standardized psychosocial stressors, including, for example, free speech or mental arithmetic tasks in front of an audience, found significantly increased epinephrine and norepinephine levels compared to healthy controls. Because elevated responsiveness of the SAM system with significantly increased levels of catecholamines was also found in other chronic inflammatory skin diseases (i.e., atopic dermatitis), it has been proposed that an overreactive SAM system may be a common feature of chronic inflammatory Th_1-mediated (psoriasis) and Th_2-mediated (AD) inflammatory skin disease. It is well accepted that the SAM system represents a major immunoregulatory system that controls various aspects of immunity. It is suggested that a dysfunctional SAM system may increase the risk of an aberrant immune response, especially under stressful conditions in which the system is activated. In fact, a distinct redistribution of leukocyte subsets or altered cytokine production after stress has been identified in psoriatics. In these studies, it was demonstrated that psychosocial stress yielded an increased number of monocytes, $CD4^+$ (Th) cells, $CD8^+$ (cytotoxic T) cells, or natural killer (NK) cells, suggesting increased mobilization of these cell types in PSO sufferers. Interestingly, the immunological changes were significantly correlated with stress-induced increases of epinephrine and norepinephrine levels in these patients. The PSO-specific immunological alterations to stress may be of clinical importance. The first leukocytic infiltrate that precedes the developing psoriatic plaque consists primarily of monocytes and CD4+ cells. After invasion into the skin, monocytes differentiate into macrophages and can act as antigen-presenting cells. Antigen presentation, however, is necessary for the initial activation of T cells. Monocytes are further able to produce chemokines that increase T-cell trafficking and infiltration into the skin. $CD4^+$ cells play a key role in the pathogenesis of PSO, and they are considered to be mainly responsible for the initiation and maintenance of the psoriatic phenotype. For example, the injection of $CD4^+$ cells into prepsoriatic skin causes the phenotypic conversion of previous symptomless skin into psoriatic skin. Activated $CD4^+$ cells (and $CD8^+$ cells) further produce pro-inflammatory cytokines such as IL-2, IFN-γ, and TNF-α, which are known to play a pivotal role in PSO pathogenesis. It should be noted, however, that, although the stress-induced accumulation of PSO-relevant immunocytes in the peripheral blood of psoriatics has been successfully documented, evidence of stress-induced infiltration of these cells into the (inflamed) skin has not been assessed and, thus, remains speculative. Animal data, however, indicate that acute stress may lead to a pronounced recruitment of immunocytes to the side of inflammation and, further, to a 200–300% higher infiltration of these cells into the (inflamed) skin. In this model, stress hormones seem to be involved because adrenalectomy abolished, and the injection of corticosterone and/or epinephrine enhanced, this effect.

Summary

In this article, important immunological and psychological factors related to the pathogenesis of PSO have been described. It has been become clear that immunological abnormalities, such as a dysfunctional Th_1/Th_2 immune response pattern, inappropriate cytokine secretion, and keratinocyte hyperproliferation, play a key role in the onset and maintenance of the disease. On the other hand, stress has been found to aggravate the skin condition in PSO. Regarding the stress–PSO relationship, it is important to keep in mind that stress can trigger the exacerbation of the disease but, in turn, may be a consequence of a worsening of skin condition and the discomfort related to an exacerbation of the disease. This may lead to a *circulus vitiousus* that is often difficult to interrupt.

In the present overview, potential psychobiological pathways, which may be involved in the close relationship of stress and the exacerbation of psoriatic skin inflammation, have been briefly discussed. In PSO patients, distinct aberrations in important immunoregulatory neuroendocrine systems such as the HPA axis and the SAM system may lead to an immunological milieu that may render (genetically disposed) individuals more vulnerable to the development and/or exacerbation of chronic inflammation. The dysfunctional responsiveness of (neuro)endocrine stress systems such as the HPA axis and the SAM system may be of specific pathological relevance when the organism is challenged by a stressor, and an adaptive regulatory control of the immune system is required.

It is of note, however, that, although this model is built on experimental evidence, several parts of it are still speculative, awaiting validation by future

research. Trying to integrate existing facts and promising working hypotheses into a model of psychoneuroimmunocutaneous communication and, further, to validate such a working model in future research is a challenge for the field of psychodermatology that may lead to new disease concepts and therapeutic regimens of PSO.

See Also the Following Article

Dermatological Conditions.

Further Reading

Darsow, U. and Ring, J. (2001). Neuroimmune interactions in the skin. *Current Opinion in Allergy and Clinical Immunology* **1**, 435–439.

Elenkov, I. J. and Chrousos, G. P. (1999). Stress hormones, Th1/Th2 patterns, pro/anti-inflammatory cytokines and susceptibility to disease. *Trends in Endocrinology and Metabolism* **19**, 359–368.

Fortune, D. G., Richards, H. L., Kirby, B., et al. (2003). Psychological distress impairs clearance of psoriasis in patients treated with photochemotherapy. *Archives of Dermatology* **139**, 458–465.

Griffith, C. E. (2003). The immunological basis of psoriasis. *Journal of the European Academy of Dermatology and Venereoloy* **2**, 1–5.

Griffith, C. E. and Richards, H. L. (2001). Psychological influences in psoriasis. *Clinical and Experimental Dermatology* **26**, 338–342.

Krueger, J. G. and Bowock, A. (2005). Psoriasis pathophysiology: current concepts of pathogenesis. *Annals of the Rheumatic Diseases* **64**, 30–36.

Picardi, A. and Abeni, D. (2001). Stressful life events and skin disease: disentangling evidence from myth. *Psychotherapy and Psychosomatics* **70**, 118–136.

Psychoanalysis

P Roazen
York University, Cambridge, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 283–286, © 2000, Elsevier Inc.

Starting Point
A Science of the Mind
Social Philosophy
Conflictedness of Life
Challenges
Conclusion

In the late 1890s, Sigmund Freud (1856–1939), who had been trained as a Viennese neurologist, created a new field, psychoanalysis, which was designed to understand and treat neurotic afflictions. An essential key to Freud's thinking lies in the character of the last days of the Hapsburg Empire. An unusual gulf between reality and official ideology stimulated a general intellectual revolt, a search for the actualities beneath the pious formulas of proclaimed truth. This revolt was led by those ideally placed to see the discrepancy because they had nothing to gain from accepting the official view, the educated Jews. Mordant irony was an essential weapon for piercing the veil of the structure of formal beliefs. The cultural conflict between East and West that had its vortex in Vienna's cosmopolitan intellectual life and the sense that liberal culture was on the verge of being undermined, were reflected throughout Freud's mature thought.

Starting Point

Freud's starting point as a therapist was the existence of inner conflicts, which interfered with the lives of suffering patients. He proposed that symptoms be looked on as substitute satisfactions, the result of a failure to deal adequately with early childhood patterns. Freud was highlighting persistent infantilism as the ultimate source of adult neurotic problems.

Freud held that neuroses are psychologically meaningful, and he interpreted them as compromise formations between repressed impulses and censoring parts of the mind. One portion of every symptom was understood as the expression of wish fulfillment, and another side represented the mental structure that reacted against the primal wish. Initially, Freud thought that neurotic anxiety arose from sexual sources; specifically, Freud indicted dammed-up sexuality as the physical basis for neurosis.

Freud conceived sexuality so broadly as to include virtually all aspects of childish pleasure-seeking. Fantasies of sexual gratification stemming from early childhood were allegedly the source of adult neurotic dilemmas. Freud proposed that a person's emotional attitude toward his or her parents encapsulated the core problems of neurosis, and he coined the term Oedipus complex to describe a boy's first childish desires for his mother and a girl's earliest affection for her father. Freud understood that someone's emotional attitude toward a family consisted of conflicting emotions involving rivalry and guilt, not just desire. Freud believed that the most troublesome feelings stemmed from emotional problems about which the individual remains unconscious.

Freud was proposing that people have motives that can be operative without their knowing anything about them. His special viewpoint was that of a psychologist, and he sought to pierce the mysteries of memory and false recollections. Freud thought that the compromise formations in constructing our image of the past were just like those in dreaming, as well as the ones underlying neurotic symptomatology and everyday slips of the tongue or pen. He thought that the past lives on in the present and that psychoanalytic treatment consisted in the exploration of each patient's early history.

Freud was ambitious as a theorist, and his notion of neurosis became part of a full-fledged system of thought. A central implication of his approach amounted to an assault on our confidence in our ability to think rationally. Freud was insisting that people are fundamentally self-deceptive. Neurosis was a form of ignorance, and Freud saw it as his task to use the power that came from enlightenment.

A Science of the Mind

Because much as Freud's work can be understood as a critique of the capacity for self-understanding, he was superlatively rationalistic about psychoanalysis itself. He thought he had discovered a science of the mind and that he had uncovered a realm of meaning that could be verified objectively. The technique of free associations, which he relied on during treatment, was one that others could be trained to use. Once patients submitted to the analytic situation, such a commitment could be used by the analyst to promote personal autonomy. Freud was relying on a structured situation in which patients saw the analyst 6 days a week for sessions lasting 50 min each for the sake of forcing people to be free, to use the terminology of Jean-Jacques Rousseau.

One of the chief defects in Freud's approach was his unwillingness to concede the philosophical underpinnings to his approach. He was convinced that psychoanalysis was capable of transforming thought and undermining previous moral positions, yet he fancied that he had been able to do so without importing any ethical baggage of his own inside his teachings. Yet Freud was clearly expressing a morality of his own; he once explained to a patient that the moral self was the conscious, the evil self being the unconscious. Freud qualified this distinction by maintaining that his approach emphasized not only evil wishes but the moral censorship that makes them unrecognizable. He was insistent that morality was self-evident at the same time that he himself supposedly held to a higher standard of ethics than humanity as a whole.

Because Freud wrote so much about abnormality, it might seem that he would have been obliged to discuss his picture of mental health, but whatever Freud had in mind has to be teased out of his system of thought because he remained loath to deal frontally with a concept such as normality. It is clear that he did not envisage a utopian version of personal happiness; anxiety and despair were to him inevitable parts of the human condition. Freud sought not to eradicate human conflicts but to teach how we can come to terms with them.

Social Philosophy

Although Freud claimed to be intent on steering away from speculative theorizing, he repeatedly allowed himself to become engaged in social philosophy. In each of the last three decades of his life he wrote a book centering on different aspects of the psychology of religion. He made the analogy between religion and obsessional neurosis and pointed out how often outer forms have obliterated the inner religious intention, as with any other self-defeating neurotic structure. The one social coercion Freud felt to be humanly unnecessary was religion. In "The Future of an Illusion" (1927), he stressed the inner instinctual core that strains beyond culture's reach. Freud's theories emphasized the divisions within the human mind; he was differing from the classical liberal tradition to the extent that he saw the individual not as a unit but as an opposed self. However, it is also the case that Freud proposed that there was a central portion of the self deep within humankind that had to remain in opposition to society.

Freud's enlightenment heritage led him to denounce religious belief in so bold a manner. As of 1927, Freud insisted that human helplessness was at the origin of religious conviction; people need religion because of the failure to outgrow the dependency of childhood. Religion is an illusion in the sense that

it is the product of wish fulfillment. Freud skeptically saw religion as a pack of lies, fairy tales that were a product of emotional insecurities. Because religion was based on irrational fears, its unreality may undermine the civilization it currently supports. Illusions are dangerous, no matter how comfortable. Freud ignored his earlier comments on religion, relating it to fears of death and guilt, and he concluded that superstition is intolerable. He was so intolerant of the infantile and the regressive that he had difficulty understanding their constructive functions.

Freud saw the family as the prototype for authority relationships. As he had argued that God the father was needed to allay the deepest fears, so he thought that the Oedipus complex also illuminated the social cohesion of political groups. Freud was suspicious of the masses and disdainful of the lower classes. His elitism lay behind a good deal of his social thinking. Religion always seemed to Freud a more intolerable irrationality than political authority. Politically, he was impressed by the extent of human inner instability and the craving for authority. Although Freud's whole form of therapy was aimed at liberation and independence, politically he was a pessimist.

Conflictedness of Life

In "Civilization and Its Discontents" (1930), Freud eloquently expressed his full sense of the conflictedness of life. He stressed the inevitable pervasiveness of suffering in civilized society. Although he could, as he had in "The Future of an Illusion," write like an eighteenth-century libertarian, here his sense of the inevitable cruelties of life was uppermost in the argument.

For civilization to be powerful enough to protect people from one another and from nature, it must, according to Freud, have at its disposal an equally intense energy. Throughout Freud's thought there is a sense of the limits of life, the truth behind the maxim that one cannot have one's cake and eat it too. Social unity can only be achieved on the ruins of human desires. People need the security of civilized life so deeply that they renounce the gratification of instincts in exchange for society. Freud concluded that the frustration of sexual and aggressive drives are entailed by their very character. Only if society can successfully internalize human instinctuality can civilization be maintained.

Challenges

Alfred Adler

As early as the years immediately before World War I, members of Freud's movement were challenging some of his central points of view. Adler (1870–1937), for example, was a socialist who went on to found a school of individual psychology apart from Freud's own. Adler had a special concern with the social and environmental factors in disease and highlighted the role of compensations for early defects in his study of organ inferiority; he was proposing that, under the best circumstances, defects in a child could create a disposition toward better performance. Adler was not as exclusively concerned with infantile sexuality as Freud was but was, instead, preoccupied with ego mechanisms and aggressive drives. In contrast to Freud's own lack of interest in politics, Adler sought to improve the world through education and psychotherapy.

Carl Gustav Jung

Jung (1875–1961) led the most painful of the secessions from psychoanalysis. Even during his period of cooperation with Freud from 1906 until 1913, Jung had hesitated to extend the concept of sexuality as broadly as Freud wished, and Jung came to interpret much so-called infantile clinical phenomena as being of secondary rather than causal importance; current difficulties, he held, could reactivate past conflicts. Jung insisted that the past can be used defensively to evade the present, a clinical point that later commanded widespread agreement. Less rationalistic and suspicious of the unconscious than Freud, Jung began to formulate his own views on the compensatory functions of symptoms. No better critique of Freud's excessive rationalism can be found than in Jung's collected works. He proposed that symptoms are always justified and serve a purpose. He was also interested in other stages of the life cycle than the oedipal one. It is still not widely known how early Jung emphasized the central importance of establishing a personal rapport between the patient and analyst if therapy is to be successful.

Further Development

The full-scale development of ego psychology was one of the main developments of psychoanalytic theory since the late 1930s, and it silently incorporated most of the central objections that people such as Adler and Jung had had to Freud's original formulations.

Following Freud's death in 1939, more attempts were made to correct the negativism that had been built into his earlier work. Freud's whole system was designed to explain motivation when a person is in conflict and the ego has relatively failed at its integrative task. As a therapist, Freud was preoccupied with pulling problems apart and tearing fixations asunder, on the assumption that the patient's

ego would be able to put the pieces back together again. For Freud, analysis was automatically synthesis; constructive processes had originally been taken for granted by him.

Freud was a master at understanding the main means of self-deception, but he ignored many processes of self-healing. Therefore, a main trend after his death was to correct this imbalance and to focus on the ego as an agency integrating inner needs and outer realities. The ego has a unifying function, ensuring coherent behavior and conduct. The job of the ego is not just the negative one of avoiding anxiety but also the positive one of maintaining effective performance. The ego's defenses may be adaptive as well as maladaptive. Adaptation is itself bedeviled by anxieties and guilts, but the ego's strength is not measured by the earlier psychoanalytic standard of what in a personality is denied or cut off but, rather, by all the extremes that an individual's ego is able to unify.

At the same time, psychology shifted from the more traditional concern with the defensive ego to the problems of growth and adaptation; it looked for the collective sources of ego development. For instance, there can be a need for a sense of identity to be confirmed by social institutions, as Erik H. Erikson (1902–1994) pointed out; and here organized religion and ritual can play a positive role. Cultural institutions not only can play a constructive role, but they help account for the variety of psychological patterns that can be found in different societies.

In his respect for the dignity of his patients, which made his innovations possible, and in his conviction that despite appearances all people are psychologically one, as well as in his individualistic goals, Freud was a great heir of the enlightenment. He was among those who are ever demanding more freedom. At the same time, however, in the development of psychoanalysis the open-ended quality of liberalism led to a revision of some of its most cherished premises; Freud represents an aspect of liberalism's self-examination. In Freud's quest for an understanding of human feelings, he transcended liberalism and joined hands with thinkers usually associated with traditions alien to it. He demonstrated the degree to which the child lives on within the adult and the way psychological uncertainties can prevent people from ruling themselves.

The humanistically oriented revisionists of Freud's views, for example, Erich Fromm (1900–1980), were trying to inject genuine humanitarianism into a psychoanalytic worldview that appeared with Freud to have ended in therapeutic despair and ethical nihilism. In correspondence and conversation, Freud acknowledged that health was only one value among others and that it could not exhaust morality as a whole. If he was wary about this whole subject of normality, it was because he realized what kind of quagmire he was in danger of entering. He touched on the subject of normalcy only on the rarest occasions. Freud typically took for granted that the people he liked best to work with were creative and self-disciplined.

Freud feared that the most original and disturbing aspects of his ideas would be destroyed by the widespread popular acceptance of his ideas in North America, yet it is questionable whether he did enough to prevent precisely this outcome. By not providing more hints about normality and not owning up publicly to the wide variety of psychological solutions he found both therapeutically tolerable and humanly desirable, Freud contributed to what he most sought to prevent. He had set out to transform Western values; he was eager to go beyond accepted good and evil. When he assaulted the maxim "love thy neighbor as thyself" as both unrealistic and undesirable, he was explicitly trying to overturn Christian ethics.

Conclusion

It is logically impossible to talk about neurosis without at the same time implying a standard of maturity as well, and yet despite how powerful psychology can be in outlining human defects and weaknesses, it has not been nearly as successful in coming to terms with the positive side of human strength and coherence. In the end, the issue of the significance of normality and its relationship to nihilism has to be left an open question. Freud's psychology did contribute to our understanding of what it can mean to be human. However, it is impossible to attempt to spell out in a definitive way the ideological implications of psychoanalysis. The writers who have been influenced by Freud constitute a wide range of people, from the most radical to some of the most conservative.

It was an old analyst and loyal disciple of Freud, Helen Deutsch (1884–1982), who had the most appropriately philosophical attitude toward the perplexing issue of normality. In her earlier years, when she was one of the most prominent teachers in the history of psychoanalysis, she made it a practice in the course of interviewing prospective analysts for acceptance into training to ask what they thought a normal person would be like. It is, of course, an ultimately unanswerable conundrum and yet one that civilized people are obliged to raise repeatedly. Like all genuine questions in social philosophy, the problem of normality can never be solved; it has to remain a real issue to the extent that we find it intolerable to contemplate a universe lacking in moral values.

See Also the Following Articles

Freud, Sigmund; Psychotherapy.

Further Reading

Ellenberger, H. F. (1970). *The discovery of the unconscious: the history and evolution of dynamic psychiatry.* New York: Basic Books.

Freud, S. (1953–1974). Civilization, its discontents (1930). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 21), pp. 59–105. London: Hogarth Press and the Institute of Psychoanalysis.

Freud, S. (1953–1974). The future of an illusion (1927). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 21), pp. 3–56. London: Hogarth Press and the Institute of Psychoanalysis.

Roazen, P. (1975). *Freud and his followers.* New York: Knopf (Reprinted 1992, New York: Da Capo.).

Psychoanalytic Theory *See:* Psychoanalysis.

Psychological Stressors, Overview

S M Monroe and G M Slavich
University of Oregon, Eugene, OR, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S M Monroe, volume 3, pp 287–293, © 2000, Elsevier Inc.

Historical and General Considerations
Conceptual Progress
Methodological Considerations and Recent Developments

Glossary

Catecholamines	Neurotransmitters, including epinephrine, norepinephrine, and dopamine, that promote sympathetic nervous system activity. They may be released in substantial quantities during stressful times.
Corticosteroids	Complex chemical compounds produced in the outer layer of the adrenal gland, which include both mineralocorticoids and glucocorticoids. Mineralocorticoids maintain salt and fluid balance in the body, while glucocorticoids have metabolic and anti-inflammatory effects and are important mediators of the stress response.
Neuroendocrine	Relating to the nervous and endocrine system, which produces endocrine secretions that help to control bodily metabolic activity.
Stress sensitization	The enhanced and progressive sensitivity of an organism to stress given repeated exposure to stressors.

Historical and General Considerations

Historical Matters

It is frequently assumed that psychological stressors are a concern of especially modern origins, or at least that they have become more prominent with recent advances in society and technology. It is also commonly assumed that with the accelerating progress of civilization, more and more people are afflicted with mental and physical disorders. Historical accounts, however, suggest that such ideas about stressors, civilization, and disease have been common for quite some time. Sir Clifford Albutt (1895: 217) expressed such sentiments quite clearly well over 100 years ago:

> To turn now... to nervous disability, to hysteria... to the frightfulness, the melancholy, the unrest due to living at a high pressure, the world of the railway, the pelting of telegrams, the strife of business ... surely, at any rate, these maladies or the causes of these maladies are more rife than they were in the days of our fathers?

The tendency to view life in stressful terms may be even more basic to human cognition than is readily

apparent. The Greek myth of Sisyphus is enlightening in this regard. The perpetual work of pushing a boulder up a mountain – only to have gravity bring it back down after each and every effort – captures some of the qualities and characteristics linked to modern views of psychological stressors. Perhaps there is something fundamental about the human condition and psyche that fosters a perception of the world as a place rife with unrelenting demands that can never be fully met, resulting in subjective states of fatigue and distress and eventually leading to ill health. Each era may bring its unique colorations to such perceptions and its own attributions regarding their origins.

It is against this psychological backdrop of belief and possible bias in thinking that modern work on psychological stressors must be examined. Psychological stressors and related concepts have been popular explanatory constructs throughout recent, and perhaps not so recent, history. As a result of their subjective allure and apparent explanatory power, these ideas have often been loosely formulated. Owing to conceptual fuzziness and ambiguity, not only has progress in science been slowed, but nonscientific issues, ideas, and biases have been permitted to masquerade as scientific truths.

The concept of psychological stressors is rich with possibilities for shedding light on important matters in adaptation, dysfunction, and disease. The concept is paralleled, though, by the potential pitfalls that accompany its intuitive appeal. The challenge is to translate the fertile ideas about psychological stressors into more precise concepts, definitions, and operational procedures. With more sound definitional and methodological procedures in place, the utility of stress concepts for understanding adaptation and maladaptation, mental and physical disorder and disease, will be better understood.

Early Ideas and Research

A broad foundation for understanding the organism's reactions to challenging environmental circumstances was laid down by Claude Bernard and Charles Darwin during the nineteenth century. Each of these two influential individuals in his own way touched on issues deriving from the tension resulting from ongoing adaptation to changing and challenging environmental circumstances. Yet it was not until the early to mid-twentieth century that such generality and complexity began to be translated into more specific terminology and technology. These efforts can be traced to at least three different lines of thought and research.

The early work of Walter Cannon dealt with ideas about common emotions and their physiological consequences, particularly with respect to the body's maintenance of homeostasis. This line of study was complemented shortly thereafter by the animal laboratory studies of Hans Selye, wherein acute and severe stressors were systematically investigated. It was in Selye's work that the concept of stress most forcefully emerged. Stress was defined as "the nonspecific response of the body to any demand" (Selye, 1976: 74). Stressors, in turn, were defined as "that which produces stress" (Selye, 1976: 78). Finally, from another vantage point, Adolph Meyer popularized the life chart methodology. This approach emphasized the importance of the dynamic interplay between biological, psychological, and social factors, such that important events within the person's biography became foci of attention for studying health and disease. Collectively, these activities, and the multiple lines of research they generated, served to initiate specific awareness of, and interest in, psychological stressors.

Arising outside of the more purposeful activities of science was another influential development that contributed to the emerging idea that psychological stressors cause both mental and physical disorders. Prior to World War II, psychopathology was commonly attributed to genetic factors or acquired biological propensities; so-called normal people devoid of such taints were thought to be largely invulnerable to mental illness. The experiences during and after World War II dramatically shifted thinking in medical and psychiatric circles to incorporate the idea that severe stress could precipitate breakdown in a previously healthy individual. Once this conceptual shift began, it underscored the multiplicity of health consequences that could be caused by severe stressors. It also opened the door for enlarging conceptual perspectives on psychological stressors by considering how less severe, yet still noxious, aspects of the social and physical environment could contribute to, or precipitate, pathology.

Conceptual Progress

Upon the foundations of stress research and theory laid down by Selye, Cannon, and Meyer, along with the influences of experiences of World War II, modern inquiry into the effects of psychological stressors became a topic of increasing interest and, eventually, of extensive empirical inquiry. Two general themes may be discerned that have underpinned advances in theory: (1) characteristics of psychological stressors and (2) individual differences in response to psychological stressors.

Stressor Characteristics

Despite general agreement about the importance of psychological stressors for health and well-being,

determining exactly what it is about stressful circumstances that is deleterious has proven challenging. An initial question of considerable theoretical importance involved the basic nature of psychological stressors: are they best viewed in a unitary manner as nonspecific demands on the organism (as postulated by Selye), or are psychological stressors more effectively viewed as a class of conditions harboring specific component characteristics of importance? Investigators from two traditions – animal and human research – have addressed this issue, with parallel and sometimes intersecting developments. Although considerable progress has been made, the general topic of elaborating stressor characteristics remains one of central importance in current thinking on psychological stressors.

Animal Laboratory Research

A great deal of work in the 1960s and 1970s was performed to determine whether specific psychological characteristics of stressors possess qualitatively distinct implications for the organism. Initially this work revealed how particular features associated with the environmental stressors might be important for predicting adverse outcomes (as opposed to the more psychologically neutral notion proposed by Selye of general or nonspecific adaptive demands). Such research went on to probe the types of psychological stressors and their effects. It became of central interest to understand in a more differentiated way the effects of diverse psychological stressors.

Animal laboratory studies adopted ingenious designs to differentiate psychological components associated with environmental stressors, with the findings from these studies demonstrating that distinctive psychological characteristics were responsible for many immediate behavioral or physiological responses. For example, specific psychological characteristics of stressors, such as undesirability or controllability, were particularly pertinent for the development of various disorders. It became clear, too, that other characteristics of stressors were important. For example, different parameters of shock administration (acute, intermittent, or chronic) produced different physiological effects in animals. Further, such differences might increase, decrease, or not influence the development of particular diseases. Finally, psychological stressors not only could influence immediate psychobiological functioning, but also could have long-term ramifications through permanent alteration of the organism.

As the importance of specificity of stressor characteristics became more apparent and accepted, questions about the specificity of stress responses also arose. What were the implications of specific stressor characteristics for different facets of psychological and physiological functioning? Such theoretical developments greatly extended the framework for inquiry, requiring attention to multiple characteristics of stressors in relation to multiple psychological and biological processes and outcomes. Relatively simple, singular response indices (e.g., corticosteroids, catecholamines) were replaced by patterns of behavioral and biological effects or profiles of neuroendocrine responses. More recently, other levels of conceptualization have been proposed. For example, psychological stressors may promote fundamental disruptions in oscillatory regulation of basic biological functions or reversions to earlier modes of functioning.

Overall, research on psychological stressors from animal research has moved beyond unidimensional and linear concepts of stressors and their effects. More recent thinking has adopted a larger framework for understanding the diverse characteristics of stressors that influence particular and varied response systems of the organism. The response systems of interest have expanded from single systems to patterns or profiles of response across multiple indices.

Human Experimental and Field Studies

Investigators of psychological stressors in humans also conducted innovative and insightful studies, both in the laboratory and in the field. Early work tended to focus on the aversive subjective attributes, particularly perception or appraisal, of psychological stressors as evaluated in an experimental setting. Yet at about this same time research on stressful life events began. It is in this area of stress research that activity on psychological stressors perhaps reached its pinnacle, in terms of both productivity and popular interest.

Extrapolating from animal laboratory studies on the one hand, and integrating with Meyer's life chart procedures on the other, Thomas Holmes and Richard Rahe first formulated the idea that distinctive changes in one's life circumstances – specific and documentable life events – could be defined and assessed in an objective manner. The work was initially based on case histories of some 5000 tuberculosis patients, from which they derived a list of 43 life events "empirically observed to occur just prior to the time of onset of disease, including, for example, marriage, trouble with the boss, jail term, death of spouse, change in sleeping habits, retirement, death in the family, and vacation" (Holmes, 1978: 46). The Schedule of Recent Experiences (SRE) was developed and published, and by 1978 more than 1000 publications had utilized this convenient method for

probing a vast range of questions pertaining to stress and illness.

The common feature associated with these disparate life changes – the stressor characteristic of primary concern – was thought to be the degree of social readjustment entailed by the event: "The relative importance of each item is determined not by the item's desirability, by the emotions associated with the item, nor by the meaning of the item for the individual; it is the amount of change that we are studying and the relationship of the amount of change to the onset of illness" (Holmes, 1978: 47). This viewpoint is consonant with Selye's ideas about stressors and stress. Hence, the psychologically neutral notion of the readjustment required of life changes was conceptualized as the characteristic responsible vulnerability to a wide variety of psychological and physical maladies.

Much as the emphasis in animal laboratory studies shifted from psychological neutral concepts of any demand, viewpoints within the stressful life events literature began to shift away from the concept of readjustment and toward emphasizing the undesirable characteristics of events. Human studies of life events consequently began to focus on the particular characteristics of psychological stressors and their potentially unique effects. The principle of specificity also was extended from the characteristics of stressors to the specific consequences of such experiences, elaborating theory about the importance of specific psychological stressors for specific responses and eventually for specific types of disorder or disease. A vast literature on this topic has appeared over the past two decades, with diverse conceptualizations of psychological stressors and myriad methods designed to measure them.

The issue of desirability of events, however, along with the more general issue involving stressor characteristics, brought into focus another important subject in the study of psychological stressors: individual differences. What might be viewed or experienced as undesirable by one person could be viewed or experienced as desirable by another. As discussed next, a variety of considerations are invoked to explain variability in effects and outcomes in relation to psychological stressors.

Individual Differences

Despite progress in conceptualizing the component characteristics of psychological stressors, and despite progress in prediction afforded by such work, considerable variability in response to psychological stressors occurs. Even under the most dire of stressful conditions, all animals or individuals do not necessarily break down. Although a refined understanding of stressor characteristics still may account for some variability in outcomes, other factors may be useful to effectively model effects of psychological stressors. Progress in understanding this issue has again come from both the basic laboratory and human studies of psychological stressors.

Animal Laboratory Research

Although there were characteristic features of physiological responses to the stressors employed in the early paradigm adopted by Selye, not all animals responded to stress in an identical manner. Further, individual differences in response were even more pronounced when the less severe types of stressors were used.

Such variables as prior experience, availability of coping responses, and other aspects of the social and experimental context were found to moderate the influence of psychological stressors. For example, when rats are exposed to electric shock, animals that cannot predict shock occurrence (via warning tones) develop a sixfold increase in gastric ulceration compared to their yoked counterparts (who receive the warning tones). Work along these lines demonstrated the delicate and often subtle interplay among stressor, social context, and resources available to the organism in determining response outcomes. These lines of study, too, suggested that individual differences in susceptibility to psychological stress could be viewed within a developmental perspective in terms of stress sensitization. Laboratory animals repeatedly exposed to severe psychological stressors can become neurobiologically sensitized to the stressors, such that relatively minor degrees of stress eventually acquire the capability of triggering pathogenic responses.

Human Life Stress Research

The importance of individual differences was perhaps more apparent in studies of human life stress and its consequences. A consistent criticism of life events research was the relatively weak association between psychological stressors and disorder. It was assumed that other considerations moderated stress effects, and the elucidation of such factors would increase the predictive capability of disorder following stressful events. Again, there were a number of factors that were believed to moderate the impact of psychological stressors, ranging from environmental factors such as social support to more individual factors such as prior experience and coping. Developmental considerations have also been important in recent theorizing about individual differences in responsivity to

psychological stressors, with the idea that prior exposure to severe psychological stressors renders the individual more susceptible to increasingly lower levels of psychological stress.

A major arena for understanding individual differences in stress susceptibility has been perception. The early and elegant laboratory studies of human stress had indicated the importance of such individual differences in perception, or appraisal, of stressors, and such thinking was readily incorporated into theory and method. Studies of life events, for example, used subjective weights of events experienced by the study participants. Once this avenue of inquiry was opened, it also brought to the forefront a variety of influences on perception, along with other factors that might influence stress responsivity. Thus, research began to focus not only on appraisal of stressors, but also on coping, social support, personality, and other considerations that in theory could moderate the effects of psychological stressors.

As research progressed along these lines, it became clear that making some of these distinctions was easier to do in theory than in method. For example, while it made good sense theoretically to consider an individual's subjective perception of psychological stressors, it was more difficult to employ such information in a scientifically sound manner. When it came to operationalization, serious problems became apparent. For example, owing to depression-based perceptual biases, a depressed person might have a skewed perception of events and rate them as particularly negative (irrespective of the objectively stressful qualities *per se*). Generally, such concerns raised an important paradox for investigations of psychological stressors. Namely, although a large part of the desired knowledge pertained to the individual's idiosyncratic appraisal of psychological stressors, methodological concerns cautioned against direct assessment of such information. Instead, alternative approaches were developed to avoid the pitfalls of using subjective reports in research on psychological stressors, as well as to avoid other problems with these methods that eventually came to light.

Methodological Considerations and Recent Developments

While concepts and methods often intertwine and, united, nurture progress, at times one or the other component may unduly influence development (for good or for bad). This situation appears applicable to research on psychological stressors, in which the methods adopted in animal laboratory research have constrained theory and methods adopted in human life stress research have misled theory on psychological stressors.

Animal Laboratory Research

The original work of Selye typically employed situations that were overpowering or unavoidable for the animal. Such conditions did not permit an evaluation of behavioral responses or other moderating influences that could influence an animal's adaptation to stressors. Further, it was realized that this methodological paradigm was not informative about psychological stressors that might be of greater ecological and evolutionary relevance (i.e., more typical with respect to the animal's natural environment and evolutionary history). Thus, such an approach readily masked the implications of less severe psychological stressors on physiology and behavior, which in turn might represent a more fertile area of inquiry into stressor effects. Finally, the nature of the stressor employed in the early stress laboratory studies also contributed to the aforementioned difficulty in differentiating physical from psychological effects, which inhibited progress in the arena of conceptual development.

Overall, the range of psychological stressors was constrained by the methods adopted. Theory, in turn, was constrained to account for the consequences of stressors under such restricted and relatively unnatural stressor conditions. More recent research has benefited from methods that encourage assessment of diverse characteristics of psychological stressors and their severities and incorporate the assessment of a wide variety of behavioral and biological response possibilities. Current perspectives based on these broader methodological approaches suggest that the organism's responses are often exquisitely specific nuances of stressors encountered.

Human Studies of Stressful Life Events

The bulk of empirical work on human life stress has been based on self-report checklist methods. The prototype of this approach is the SRE, the instrument that catalyzed research in this area. The popularity of the SRE was likely due to the combination of the intuitive appeal of the stress concept, the apparent objectivity of the method, and the overall impression of scientific legitimacy.

The methodological paradigm launched by the SRE, however, embodied several problems. It became clear that subjects did not report life events in a reliable manner over time and that investigators did not adequately control for the directionality of effects in research designs (e.g., being depressed initially could bring about life events such as trouble at work,

difficulties with spouse, and so on). Indeed, many of the initial items on the SRE were direct indicators of disorder or illness. For example, some of the key criteria for defining clinical depression were represented in the original SRE (e.g., major change in eating habits, major change in sleeping habits). If measures of life events were directly confounded with the presence of disorder or were contaminated by the effects of pre- or coexisting disorder, then clearly general theory about psychological stressors, as well as theory about the characteristics of psychological stressors, rested on flawed information.

In response to these methodological concerns, other investigators designed semistructured interview protocols and developed explicit guidelines, decision rules, and operational criteria for defining and rating life events. These developments further highlighted serious problems with self-report checklist methods. For example, there was too much subjective leeway permitted in defining what constitutes an event with self-report procedures, resulting in considerable variability of content within ostensibly uniform categories of events. In order to have a more firm methodological foundation, the more elaborate and extensive interview and rater-based procedures were employed, which helped to standardize measurement across individuals.

In general, such interview and rater-based approaches have been found to enhance the reliability of life event assessments and to provide stronger predictions of particular kinds of disorders following the occurrence of psychological stressors. Procedures such as these also provide a solid foundation upon which to build in terms of developing taxonomies of psychological stressors and their effects. Although such approaches are more time- and labor-intensive to implement, they represent the current-day gold standard for assessing psychological stressors.

Human Studies Employing Other Measures of Psychological Stressors

There have been other methods in which psychological stressors have been defined and studied. None of these approaches has received the degree of attention devoted to the work on life events, yet each may have useful properties for the study of psychological stressors. Two lines of investigation are noteworthy.

Many investigations have targeted people who experience a specific life event and compared these individuals with controls who do not experience the event. For example, individuals who become unemployed are compared to individuals who do not experience this event in relation to a variety of psychological and physical processes and outcomes. Such work is useful for examining a potentially more homogenous process with more readily identifiable outcomes. On the other hand, such studies may oversimplify the psychological stressors associated with an event and not specifically articulate the different components within the general event that are most pernicious for health and well-being. For example, the effects can be partitioned into a variety of stressful themes that, although often intercorrelated, may not have uniform effects. Thus, although people who become unemployed in general may experience a loss of self-esteem, loss of income, loss of daily schedule, and so on, each particular situation may pull more or less for heightened responses along these different dimensions. Work sensitive to such variability in the component characteristics will be most useful for research on psychological stressors.

Finally, there also have been efforts to measure psychological stressors through questionnaire or diary methods, inquiring about minor but common daily events, chronic conditions, appraisal processes, and other indicators or correlates of psychological stressors. A promising recent avenue of research involves ecological momentary assessment, where subjects can be prompted throughout the day to respond to queries about their circumstances and psychological states. Such procedures help minimize problems with standard retrospective methods.

In closing, it is appropriate to return to the concerns and caveat with which the article began. The specter of possible biases in the measurement of psychological stressors must be consistently borne in mind, and methods employed must be rigorously attentive to such concerns, to provide a solid empirical foundation upon which theory and research can build for this important area of investigation.

See Also the Following Articles

Affective Disorders; Animal Models (Nonprimate) for Human Stress; Cardiovascular System and Stress; Cognition and Stress; Control and Stress; Coping Skills; Corticosteroids and Stress; Depression Models; Environmental Factors; Immune Suppression; Life Events Scale; Psychosocial Factors and Stress; Selye, Hans; Social Support; Stress, Definitions and Concepts of; Stress Generation.

Further Reading

Allbutt, C. (1895). Nervous diseases and modern life. *Contemporary Review* **67**, 217.

Brown, G. W. and Harris, T. O. (1989). *Life events and illness*. London: Guilford Press.

Cohen, S., Kessler, R. C. and Gordon, L. U. (eds.) (1995). *Measuring stress: a guide for health and social scientists*. New York: Oxford University Press.

Dohrenwend, B. P. (ed.) (1998). *Adversity, stress, and psychopathology.* New York: Oxford University Press.
Holmes, T. H. and Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research* **11**, 213–218.
Lazarus, R. S. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.
Monroe, S. M. and Harkness, K. L. (2005). Life stress, the 'kindling' hypothesis, and the recurrence of depression: considerations from a life stress perspective. *Psychological Review* **112**, 417–445.
Post, R. M. (1992). Transduction of psychosocial stress into the neurobiology of recurrent affective disorder. *American Journal of Psychiatry* **149**, 999–1010.
Selye, H. (1976). *The stress of life* (2nd edn.). New York: McGraw-Hill.
Stone, A. A., Shiffman, S. S. and DeVries, M. W. (1999). Ecological momentary assessment. In: Kahneman, D., Diener, D. & Schwarz, N. (eds.) *Well-being: the foundations of hedonic psychology.* New York: Russell Sage Foundation.
Weiner, H. (1992). *Perturbing the organism: the biology of stressful experience.* Chicago, IL: University of Chicago Press.
Weiss, J. M. (1972). Psychological factors in stress and disease. *Scientific American* **226**, 104–113.

Psychoneuroimmunology

R Dantzer
Integrative Neurobiology, Bordeaux, France

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Dantzer, volume 3, pp 294–298, © 2000, Elsevier Inc.

Immune System
Relationships between the Nervous System and the Immune System
Functional Consequences of Relationships between the Nervous System and the Immune System
Immune Influences on Brain Functions
Implications for Psychosomatics

Glossary

Cytokines	Secreted regulatory proteins in response to external insults that control the survival, growth, differentiation, and effector function of tissue cells.
Immune system	A complex network of specialized cells and organs (lymphoid organs) that distinguishes between self and nonself and defends the body against infections from agents such as bacteria, viruses, fungi, and parasites.

Immune System

Immunology deals with understanding how the body distinguishes between what is self and what is nonself. The classic hallmark of the immune system is its ability to recognize and kill foreign substances and at the same time recognize but not destroy normal host tissues. To protect against potentially pathogenic microorganisms that are foreign to the body (antigens), the host uses white blood cells, or leukocytes, that originate from precursors in the bone narrow. These cells circulate through the blood and lymph system or, for some of them, are fixed in specific tissues. Innate immunity represents the first line of defense against antigens and involves phagocytic cells that are mononuclear (e.g., macrophages) and polymorphonuclear (e.g., neutrophils). These cells capture microbes and digest them, and the mononuclear macrophages also present antigen fragments to lymphocytes. Lymphocytes have highly diverse specific receptors, which allow them to recognize antigens. Adaptive immunity is based on the proliferation and differentiation of antigen-specific lymphocytes into effector cells. The T lymphocytes, so designated because they differentiate and mature in the thymus gland before traveling to the spleen and lymph nodes, are the first of the lymphocytes to be activated. Following contact with antigen-presenting cells such as macrophages, lymphocytes proliferate and destroy the infected cells, in which case they are called cytotoxic T lymphocytes, or they help other lymphocytes to intervene.

There are many different subclasses of T cells that can be distinguished from each other by specialized cell surface molecules. For instance, cytotoxic T lymphocytes are called CD8 positive T cells because they express a glycoprotein known as CD8. CD4 positive T cells, which are markedly depressed in patients suffering from acquired immunodeficiency syndrome, are essential for the activation of macrophages and

B lymphocytes. Other lymphoid cells, called natural killer (NK) cells, are specialized in the recognition of tumor cells. B lymphocytes, so designated because they develop in the bone marrow of adults, differentiate into plasma cells that produce antibodies when they bind antigen and are co-stimulated by a subset of T cells. Antibodies, also called immunoglobulins (Ig), are soluble molecules that attach specifically to antigens, making their destruction possible through a series of complex reactions involving complement molecules. An immune response results in the generation of an increased number of differentiated memory lymphocytes that, upon re-exposure to the same antigen, generates more rapid and effective response. Some antigens induce the production of specialized antibodies, known as immunoglobulins E (IgE), that bind to the surface of mast cells. Recognition of an antigen or allergen by IgE attached in this manner induces the release of various mediators (e.g., histamine and serotonin) that are contained in the secretory granules of mast cells and are responsible for an allergic reaction.

The various cells that are involved in immune reactions communicate with each other by cell-to-cell contact via specialized adhesion molecules and by soluble mediators known as cytokines or interleukins. These two forms of cell-to-cell communication are essential for the initiation and regulation of the host response to infection. Each category of immune cells produces characteristic cytokines. Activated monocytes and macrophages produce the proinflammatory cytokines that are necessary for the development and regulation of the inflammatory response and include interleukin-1 (IL-1), IL-6, and tumor necrosis factor-α (TNF-α). Two subsets of CD4 positive T cells, or T helper (Th) cells, can be distinguished on the basis of their profile of cytokine secretion and the type of immune response they promote. Th 1 cells secrete IL-2, interferon-γ (IFN-γ), and TFN-α, whereas Th2 cells release IL-4, IL-6, and IL-10. Th1 cells are major players in delayed-type hypersensitivity and proinflammatory responses, whereas Th2 cells induce B cell growth and differentiation.

Relationships between the Nervous System and the Immune System

Although the immune system has long been viewed as a self-regulated system, there is now a wealth of evidence in favor of the existence of intricate anatomical and functional relationships within the nervous system. In particular, branches of the autonomic nervous system innervate the spleen and lymph nodes. Immune cells have functional membrane receptors for both classical neurotransmitters and neuropeptides. This is important because it implies that immune cells proliferate and differentiate in a microenvironment, the composition of which is dependent on the activity of the sympathetic nervous system. In addition, leukocytes themselves are able to produce and release many peptides that have been identified initially in the neuroendocrine and nervous system, such as proopiomelanocortin, corticotropin-releasing hormone, growth hormone, and insulin-like growth factors. Conversely, microglial cells, which are the equivalent of macrophages within the central nervous system, are able to synthesize and release proinflammatory cytokines such as IL-1 and TNF-α that are involved in the regulation of the inflammatory response at the periphery, and brain cells also express receptors for these and other cytokines.

Functional Consequences of Relationships between the Nervous System and the Immune System

The commonality of messengers and receptors between the immune system and the nervous system provides new vistas into the old notion that psychological factors can influence the development and evolution of somatic diseases. Objective studies on the influence of psychosocial factors and specific emotional states on immune responses are now possible thanks to the development of quantitative methods for assessing changes in immunity in animals and human subjects. In animal studies, exposure to a variety of stressors, including painful electric shocks, physical restraint, social separation, and the odor of another animal that is stressed, can depress NK cell activity and T and B lymphocyte proliferating response to nonspecific mitogens. Psychological aspects of stressors, particularly the subject's ability to predict and control their occurrence, are more important than their severity, frequency of occurrence, and duration. The effects of stress on immunity depend on the immune response under consideration, and a given stressor can either suppress or enhance different immune responses in the same individual. For example, mice that were exposed for 24 h to the odor of another stressed mouse display a decrease in IL-2 production by splenocytes stimulated with a T cell mitogen, a reduction in NK cell activity, and an increase in IgM and IgG antibody titers against a T cell-dependent antigen. These apparently disparate results can be interpreted as an effect of stress on the Th1/Th2 balance, in the form of activation of Th2 cells and downregulation of Th1 cells. Human studies show that immune functions can be altered by various experimental procedures, such as emotion induction under controlled laboratory conditions

and naturally occurring life events, such as academic examination, difficulties in marital life, and intractable diseases (e.g., cancer, early dementia) affecting the spouse. Major depression is accompanied by alterations in the distribution of T cell subsets and by declines in nonspecific measures of immunity such as NK cell activity and mitogen-induced lymphocyte proliferation. However, there are also signs of immune activation, in the form of increased numbers of leukocytes, neutrophils, and activated T lymphocytes, and an increased production of Th1-like cytokines.

Pathways involved in the mediation of the effects of stress on immunity can be identified by interfering with the activity of the autonomic nervous system and endocrine glands that are known to be involved in neuro-immune interactions. Experimental studies have made it clear that glucocorticoids are not the sole mediator of the effects of stress on immunity. Activation of the autonomic nervous system plays an important role, probably by altering the molecular composition of the microenvironment in which lymphocytes proliferate and differentiate.

The possible consequences of stress-induced immune alterations on the host resistance to infectious disease or the progression of tumor cells are not easy to assess. The relationships between stress and disease are already extraordinarily complex, and in most cases, the importance of the association between stress and illness behavior (symptom awareness and use of health services) is blurring the picture. It is now obvious that there is no global answer to this issue. In each case, a slow and cautious approach, in which the whole set of relationships among a stressor, a specific pathway, immunity, and disease is examined systemically, is more likely to yield valid answers.

Immune Influences on Brain Functions

Infection with foreign agents, in the form of infectious microorganisms or nonreplicating antigens, has long been known to cause a classical stress response, involving activation of the hypothalamic-pituitary-adrenocortical (HPA) axis and increases in catecholamine turnover in the brain and periphery. It is now apparent that these neuroendocrine and neurochemical changes form part of a larger, physiological regulatory system that modulates functional activities of lymphocytes and phagocytes. Molecules that signal the brain that an immune event has been initiated in the periphery have been identified. They are represented by proinflammatory cytokines, such as IL-1 and TNF-α, that are synthesized and released by activated macrophages and monocytes during the course of an infection. These peripherally released cytokines act on the brain by two pathways: a neural pathway that involves sensory fibers innervating the bodily site of infection and a humoral pathway that involves blood-borne cytokines and activated circulating immune cells. In the first case, the transmission of the immune message to the brain is direct, whereas in the second case, it needs to be relayed by specialized parts of the brain that have a deficient blood–brain barrier and are known as circumventricular organs because of their anatomical location around the brain ventricles. In response to these peripheral signals, glial cells produce proinflammatory cytokines that, by acting on neural target cells, activate the HPA axis, resulting in an increase in circulating glucocorticoids. This response is important for the regulation of the immune response, as glucocorticoids feed back on activated immune cells and attenuate the synthesis of proinflammatory cytokines. Glucocorticoids also act on target cells of proinflammatory cytokines to block the transcription effects of these cytokines. The lack of a glucocorticoid response to an inflammatory stimulus is responsible for development of the disease in experimental models of autoimmune diseases in animals, whether the disease is induced by the administration of specific antigens (e.g., adjuvant-induced arthritis and experimental allergic encephalomyelitis in the Lewis rat) or develops spontaneously (e.g., autoimmune thyroiditis in the obese chicken). In addition to their potent effects on the HPA axis, proinflammatory cytokines also induce fever and nonspecific behavioral responses to infection, such as malaise, pain, fatigue, sleepiness, anorexia, apathy, social withdrawal, irritability, and cognitive deficits. The development and duration of these physiological and behavioral alterations are regulated by molecules called cryogens that oppose the effects of proinflammatory cytokines. Cryogens include neuropeptides, such as vasopressin and α-melanotropin, and anti-inflammatory cytokines, such as the IL-1 receptor antagonist (IL-1RA) and IL-10.

The profound influences of cytokines on brain functions offer insights into understanding why patients afflicted with cancer or hepatitis C experience flu-like symptoms and develop acute psychotic or depressive episodes when injected with recombinant cytokines. Furthermore, many chronic inflammatory conditions have been demonstrated to be also associated with priming of the brain cytokine compartment, so that a local or a systemic immune challenge can precipitate the occurrence of cognitive alterations and mood disorders. This is particularly the case in

aged persons, obese subjects, and patients suffering from atherosclerosis. It could explain the higher prevalence of mood and mild cognitive disorders that is associated with such conditions. In addition to peripheral influences on brain cytokines, it is important to note that proinflammatory cytokines are also induced in the brain by infiltrating blood cells after a brain trauma, a stroke, or a local immune mediated inflammatory response that compromises the blood–brain barrier. The proinflammatory cytokines that are produced in these conditions play a pivotal role in the pathogenesis of the corresponding neurological disease. They induce proliferation of astrocytes and neurodegeneration via necrotic and apoptotic routes, and they are also responsible for the psychopathological and neurological alterations accompanying the disease (e.g., AIDS-related dementia).

Implications for Psychosomatics

Demonstration of the potent behavioral and psychological effects of cytokines has important implications for understanding the relationships between stress and disease. It is no longer tenable to go only one way and claim that life events interacting with personality traits, mood, and psychological variables play a causal role in the pathogenesis of immune-related diseases because of their autonomic and hormonal concomitants. Emotional dispositions and psychological characteristics can be the result of alterations in brain functions caused by still undetectable alterations in immune functions. In the case of cancer, for example, feelings of helplessness and hopelessness that have been related repeatedly to the onset and progression of a tumor may be secondary to central effects of soluble products released by immune cells of the host or even tumor cells during the early stage of the disease, when cancer is still not clinically detectable. The inflammatory response to cancer treatments, especially radiotherapy; can have similar consequences on brain functions.

Research on the effects of psychosocial factors on the onset and progression of diseases ranging from AIDS to allergies and musculoskeletal disorders is now taking a new turn. In an organism in which physiological systems are interconnected by reciprocal regulatory systems, a dysfunction in one system is likely to affect the functioning of other systems. In addition, the dysfunction can be at the level of the regulatory process itself rather than within one of the regulated organs. The immune system has long been represented as a self-regulated system that is relatively insensitive to the host's physiological turmoil. It is now clear that this view is inappropriate, and that reciprocal connections between the brain and the immune system play an important part in the coordinated response to infection and inflammation. The challenge now is to understand how such regulatory, homeostatic mechanisms can be compromised and lead to disease.

See Also the Following Articles

Cytokines; Cytotoxic Lymphocytes; Disease, Stress Induced; Health Behavior and Stress; Hypothalamic-Pituitary-Adrenal; Immunity; Psychosocial Factors and Stress.

Further Reading

Ader, R., et al. (eds.) (2006). *Psychoneuroimmunology* (4th edn.) New York: Elsevier.

Konsman, J. P., Parnet, P. and Dantzer, R. (2002). Cytokines and sickness behavior: mechanisms and implications. *Trends in Neurosciences* **25**, 154–159.

Perry, V. H. (2004). The influence of systemic inflammation on inflammation in the brain: implications for chronic neurodegenerative disease. *Brain, Behavior, and Immunity* **18**, 407–413.

Segerstrom, S. C. (2005). Optimism and immunity: do positive thoughts always lead to positive effects? *Brain, Behavior, and Immunity* **19**, 195–200.

Psychosocial Factors and Stress

J Siegrist
University of Duesseldorf, Duesseldorf, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Siegrist, volume 3, pp 299–303, © 2000, Elsevier Inc.

General Background
Overview of Psychosocial Factors Influencing Stress and Health
Implications for Intervention

Glossary

Self-efficacy	Personal beliefs about one's ability to master challenges.
Self-esteem	Firm feelings and beliefs about one's own worth as a person (influenced strongly by experience on how significant others react toward one's self early in life).
Social anomie	Lack of social rules and orientations in a social group, often associated with feelings of anxiety or helplessness.
Social capital	Patterns of cooperative social exchange in neighborhood and community life characterized by mutual commitment and trust.
Social role	Set of expectations or norms (duties, options) directed toward people who hold important social positions (e.g., family role, work role).
Type A behavior	Complex of behaviors and emotions defined by excessive striving, competitiveness, hostility, and impatience.

General Background

In evolutionary terms, the development of the human brain, particularly its neocortical structures, has critically advanced the species' capacity to cope with environmental challenges and threats. As a consequence, inherited biobehavioral patterns of response to threats to the integrity, survival, and reproduction of the organism have become more variable and complex compared to those of the nonhuman primates. Increasingly, these patterns of response depend on the appraisal of stimulus properties and available opportunities for successful agency and are modulated by prior experience, learned skills, and interpersonal support. This may hold true for responses to natural and man-made disasters and for chronically demanding socioenvironmental conditions. However, in all these instances, a person's sense of the control over the challenge and over his or her actions in dealing with it must be considered a crucial variable that triggers the quality and intensity of stressful experience. Stressful experience is defined as the experience of a mismatch between the demands put on an individual in a challenging situation and his or her ability to meet the situation with adequate responses. At the level of emotional reactions, this mismatch is likely to elicit feelings of anxiety, irritation, anger, disappointment, or even helplessness and despair. These emotions are often paralleled by sustained neuronal and neuroendocrine activation due to the existence of powerful neuronal circuits that link neocortical information processing with mesolimbic and brain stem activity.

Humans are particularly vulnerable to the distress-enhancing effects of socioenvironmental challenges and threats because the preservation of a person's self relies so heavily on his or her social surrounding. In human infancy, attachment behavior, which develops through reciprocal exchange with a caregiver, is essential for survival and normal psychological functioning. Later in childhood and adolescence, a person's social identity emerges from recurrent exchange with significant others. Moreover, in adult life, core social roles, such as the work role and the family and marital role, serve to link the individual person with a structured, goal-oriented social environment. Thus, important life domains that matter most for the individual person are inherently social in a cooperative or competitive sense.

Placed in this context, human health and well-being are closely related to successful self-regulation in a conducive social environment, and a stressful experience is likely to result from failed opportunities for successful agency, goal achievement, and reward experience in interpersonal exchange. The question then is what analytical concepts have been developed by the social and behavioral sciences to identify relevant socioenvironmental and personal triggers of recurrent stressful experience. The term psychosocial factor is used as a summary label to characterize the still-heterogeneous state of the art in this field. In other words, scientific progress has not yet resulted in a universally accepted taxonomy of stress-eliciting or stress-buffering social and personal (psychological) conditions.

This article presents a brief overview of some of the major concepts of identifying psychosocial influences

on stress and health. The two main procedures mediating this association are (1) enhanced neuronal/neuroendocrine/neuroimmune activation and (2) initiation of health-damaging behaviors.

There are multiple ways in which social and personal factors interact in influencing health and disease. It is useful to think of a stream of events where the most distal parts, the upstream factors, relate to the society at large or its main institutions (the macrosocial influences) and where the most proximal parts, the downstream factors, concern an individual's close or immediate social surroundings (the microsocial influences) and his or her personal ways of coping with the environment. In the following section, the discussion of psychosocial factors starts with the more distal parts of this stream and then proceeds to more proximal phenomena, including person characteristics. Finally, special attention is given to concepts that focus explicitly on the interaction between person and environment.

Overview of Psychosocial Factors Influencing Stress and Health

Macrosocial and Microsocial Factors

Every human society is characterized by a set of social values, norms, and institutions that are instrumental for the survival and growth of its members. If these values and norms lose their validity and meaning, or if society's patterns of social exchange become unpredictable and unstable, individuals tend to suffer from states of social anomie (lack of rules and orientations). A number of epidemiological investigations have explored the adverse effects on health produced by sociocultural instability, rapid social change, or a high level of social anomie. A majority of these studies found evidence of elevated risks of subsequent physical and mental illness.

One type of study was conducted in native societies that, after a long period of social stability, experienced rapid sociocultural transformation. As a consequence, less solidarity was witnessed among group members, and more deviant behavior, including violence, crime, and disruption of social ties, occurred. Equally distressing, the internalized cultural canon, which had been developed over many generations, broke down. Essential beliefs and ways of interpreting life were lost or fragmented. Under these conditions, cardiovascular risk factors, particularly hypertension, and coronary heart disease manifestation became more prevalent.

Similar observations were made by investigators who followed migrant populations from their original places to new sociocultural environments. Interestingly, migrants who were able to maintain strong social ties with their traditional groups in the new environment were more likely to be protected against subsequent health risks compared to migrants who deprived themselves from these traditional bonds. This observation led to the conclusion that a stable social network and a particular quality of social exchange, termed social support, may act as a protective resource in coping with stressful circumstances. Social support is defined as the experience of, or access to, social relationships that offer mutual understanding and trust and that recurrently elicit positive emotions.

There is now convincing evidence that negative health effects result from social separation and social isolation, either due to loss of a loved one or due to poor opportunities for developing or maintaining relationships. Again, under these conditions, social support is less likely to be experienced and thus a strong psychosocial resource against stress is lacking.

While close social ties and social support denote downstream phenomena close to the intimate experiences made by individual persons, the more distal, macrostructural conditions of participating in the social fabric were shown to affect health as well. In particular, disinvestment in social capital was found to exert strong effects on morbidity and mortality at the level of whole populations. The term disinvestment in social capital describes a decline in reciprocal social exchange in neighborhood and community life, including a lack of solidarity and trust. For instance, when an area has a higher level of poverty, less investments are made into shared resources of community life, and hostile encounters or competitive struggles are more frequent.

Disinvestment in social capital is one of several health-adverse developments resulting from widening social inequalities. The demonstration of an inverse social gradient of morbidity and mortality within and between populations in a large number of advanced societies must be considered one of the most consistent and important findings of modern social epidemiology: the lower a person's socioeconomic status (in terms of educational attainment, income, and occupational standing), the higher his or her risk of poor health. These differentials are substantial. For instance, mean life expectancy in countries such as Finland or the United Kingdom is shortened by 5 to 6 years among members of the lowest as compared to the highest socioeconomic status group. It is important to note that these social differentials in health are spread across the whole of the society rather than being concentrated on a deprived group at the bottom of the society. So what are the causes of relative social deprivation in health?

Research repeatedly came to the conclusion that access to or quality of health care is not the main cause, as might be expected. Rather, unhealthy behaviors and exposure to stressful material and psychosocial environments, including lack of appropriate resources, must be considered the main determinants of social inequalities in health in advanced societies. These latter conditions include poor living and working conditions, threats to, or experience of, unemployment and income loss, and unstable social networks. Under these circumstances of relative social deprivation, feelings of anger, disapproval, and hopelessness are more frequent, intense, and long-lasting, thus aggravating the burden of stressful experience (see later).

Person Characteristics

Clearly, not every individual reacts to stressful environments in the same way. On the one hand, specific genetic or psychosocial factors may protect a person from adverse effects on health even under highly demanding circumstances. On the other hand, a person's threshold of susceptibility may be very low, thus predisposing him or her to illness onset under conditions of moderate stress. With the introduction of methods of molecular medicine into epidemiological studies, important advances in identifying people at elevated risk of disease were achieved. For instance, combining sociological information on stressful life events with genetic information on specific polymorphisms of the serotonin transporter receptor gene was shown to improve the prediction of early onset of depression in young adults.

One direction of research has focused on psychological traits that predispose an individual to spend high levels of energy. This pattern is functional if demands or threats are faced. However, in less demanding contexts, this may harm the organism by recurrent excessive physiological activation. Originally, researchers were interested in so-called type A behavior, a complex of behaviors and emotions defined by excessive striving, competitiveness, hostility, and impatience. Type A individuals were identified on the basis of specific psychomotor and speech characteristics, such as unrest, hectic, inability to relax, loud speech, and intrusive behaviors. As can be expected from a stress physiological perspective, excessive sympathoadrenergic arousal predisposes these individuals to cardiovascular dysfunction and disease.

Today, however, refined concepts following this direction of research are more successful in predicting future cardiovascular risk. One such concept concentrates on cognitive and affective components of hostility. The former include negative beliefs about people and negative attitudes toward others, such as cynicism and mistrust, whereas the latter focuses on a frequent and intense experience of anger. A high level of hostility was found to promote health-adverse behaviors such as smoking and alcohol use in adolescence and young adulthood. Moreover, increased sympathetic function in combination with decreased parasympathetic function was observed. This observation could explain at least part of the direct statistical associations between hostility and coronary heart disease incidence that have been reported from several studies.

A similar approach emphasizes cognitive-motivational mechanisms of excessive energy spending. People characterized by overcommitment when facing demands tend to misjudge (e.g., underestimate) challenging stimuli and to expose themselves to multiple obligations. This may be due to their strong desire to control or dominate their environment, to exceed others, and to experience approval. For instance, workaholic behavior in occupational contexts can be interpreted in this context (see later). After a period of continued overcommitment, these individuals are susceptible to a state of vital exhaustion and psychophysiological breakdown, particularly when confronted with loss of control and failure. Overcommitment has been identified as a cardiovascular risk factor in epidemiological and experimental investigations.

A second line of research demonstrates adverse health effects of passive behaviors and associated low levels of energy spending. Importantly, inhibition or withdrawal of activity due to depressive mood and cognitions of helplessness and lack of self-esteem was found not only in reaction to but also preceding any experience of severe threat or loss. Some investigations conclude that individuals characterized by these psychological traits are at elevated risk of suffering from reduced immunocompetence and its pathophysiological consequences. Thus, over- and underactivation of a person's potential of vital energy in relation to self-regulating processes affect health at more fundamental, cellular and subcellular, levels, e.g., by increasing endogeneous oxidation and decreasing DNA repair capacity.

Although these behavioral patterns are influenced, to some extent, by genetic predisposition, evidence indicates that socioenvironmental factors have a strong impact on gene expression via methylation of DNA. Again, combining research on epigenetic processes with sociobehavioral research has promise for advancing the frontiers of knowledge.

In conclusion, distinct personal traits have been identified that must be considered psychological risk factors or protective factors. Protective factors

include a sense of mastery and self-efficacy (i.e., positive beliefs about one's agency) and favorable self-esteem, often paralleled by optimism and trust. Risk factors, as evidenced earlier, include overactive behaviors such as hostility, overcommitment, and excessive striving. These traits carry the risk of precipitating feelings of frustration and helplessness and triggering states of vital exhaustion. Moreover, risk factors include those cognitive, motivational, and affective traits that predispose individuals to withdraw from activity, to experience low levels of personal control and self-esteem, and to suffer from depressive mood and negative attributional style.

Person–Environment Interaction

Traditionally, the disciplines of medical sociology and social epidemiology have emphasized the role of socioenvironmental factors in health and disease, whereas medical psychology and psychosomatic medicine have stressed the importance of personality factors. Therefore, despite the obvious need for a more comprehensive analysis, few attempts were made to combine the two disciplinary approaches. Of these attempts, two developments deserve special attention.

First, in a period of declining impact of psychoanalytic thinking on etiological research, new interest emerged in understanding how socioenvironmental and personal factors interact in promoting health and triggering disease in a life course perspective. For instance, birth cohort studies demonstrate that adverse social circumstances of parents at the time of their children's birth increase the risk of premature mortality in children's later life. Explanations for this refer to elevated vulnerability during pregnancy and experience of cumulative cognitive, behavioral, and social disadvantage during childhood.

Among early psychological risk factors, a disturbed pattern of attachment between mother and child in infancy was shown to be of particular significance for later well-being. However, favorable social circumstances in childhood can compensate for these deficits to some extent by promoting children's resilience.

A second line of extensive research on how person and environment interact in producing stress-related disease has been concerned with the impact of working life on health. In contemporary societies, the work role still defines one of the crucial social roles in adulthood with powerful effects on successful or unsuccessful self-regulation. The adverse effects of job loss and related social isolation on health were mentioned earlier. Additional findings demonstrate that among low-status members of the workforce, stressful, health-damaging jobs in terms of a combination of high demands and low control are more prevalent. In this context, the model of effort–reward imbalance at work is of interest, as it explicitly combines situational and person characteristics.

This model claims that lack of reciprocity between costs spent and gains obtained at work (high-cost/low-gain conditions) defines a state of emotional distress with special propensity to autonomic arousal and neuroendocrine stress response. Effort at work is spent as part of a socially organized exchange process to which society contributes in terms of rewards (money, esteem, and career opportunities). An imbalance between effort and reward is maintained if no alternative choice is available in the labor market. Moreover, people characterized by a high level of overcommitment and related cognitive distortion of demand appraisal (see earlier discussion) are equally prone to experience an effort–reward imbalance. Prospective epidemiological studies documented that stressful work in terms of high effort, low reward, and low control doubles the risk of incident fatal or nonfatal cardiovascular disease. In addition, onset of depression is more likely to occur.

Implications for Intervention

This article documents a solid body of research on the role of distinct psychosocial factors in triggering stress-related disease (risk factors) and in promoting health (protective factors). It highlights that both upstream (macrosocial) and downstream (microsocial and personal) factors are important in explaining elevated risk of morbidity and mortality. Finally, a combined approach focusing on person–environment interaction, including analyses of epigenetic processes, holds special promise for advancing scientific understanding and its consequences for prevention.

At all levels discussed so far, implications for intervention are obvious. At the macrostructural level, however, it is difficult to imagine how a reduction in the amount of relative social deprivation can be achieved. Policy measures such as limitations in income inequality, labor market regulations, and allocation of means to improve social capital can hardly be realized on the basis of anticipated health gain.

A more realistic approach toward intervention concerns the mesosocial level. For instance, several theory-based concepts of workplace health promotion have been implemented and evaluated. Reductions in blood pressure, heart rate variability, and levels of stress hormones and atherogenic lipids are examples of favorable health outcomes following task reorganization with enlarged decision latitude, improved social support, and improved leadership skills of superiors.

Similarly, more comprehensive intervention programs at the level of community, neighborhood, or school are being developed. Some of these programs pay special attention to sustained health gain in socially deprived groups or in high-risk groups. Examples are early-childhood family education programs, preschool programs, and adolescent health promotion activities. Improving successful parenting in socially deprived families was found to strengthen children's cognitive and emotional development.

Finally, an obvious level of intervention concerns the microsocial and personal level. Several behavioral intervention techniques have been developed and tested in both primary and secondary prevention. These include stress management, promotion of self-control and self-efficacy, personal empowerment, and an increase in social skills and coping effectiveness. These techniques proved to be particularly successful in reducing health-adverse behaviors such as unhealthy diet, smoking, alcohol, and drug consumption. In addition, the provision of social support and the formation of self-help groups are instrumental in reinforcing newly acquired attitudes, motivations, and behaviors. In summary, a rich potential of interpersonal and behavioral intervention approaches is now available that supplements traditional biomedical treatment techniques and that provides new opportunities for primary and secondary prevention. It is important to emphasize that this potential is based on scientific evidence obtained from pioneering research on psychosocial factors, stress, and health.

See Also the Following Articles

Chronic Social Stress: GR Sensitivity in Leukocytes; Hostility; Life Events and Health; Self-Esteem, Stress and Emotion; Social Capital; Social Networks and Social Isolation; Social Support; Type A Personality, Type B Personality.

Further Reading

Antoniou, A. S. G. and Cooper, C. L. (eds.) (2005). *Research companion to organizational health psychology.* Cheltenham, UK: Edward Elgar.

Bandura, A. (1997). *Self efficacy: the exercise of control.* New York: Freeman.

Berkman, L. and Kawachi, I. (2000). *Social epidemiology.* Oxford, UK: Oxford University Press.

Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression. Moderation by a polymorphism in the 5-HTT-gene. *Science* **301**, 386–389.

Marmot, M. and Wilkinson, R. (eds.) (2005). *Social determinants of health*. Oxford, UK: Oxford University Press.

Orth-Gomer, K. and Schneiderman, N. (eds.) (1996). *Behavioral medicine approaches to cardiovascular disease prevention.* Mahwah, NJ: Erlbaum.

Siegrist, J. and Marmot, M. (eds.) (2006). *Social inequalities in health: new evidence and policy implications.* Oxford, UK: Oxford University Press.

Turner, J. R., Cardon, L. R. and Hewitt, J. K. (eds.) (1995). *Behavior genetic approaches in behavioral medicine.* New York: Plenum Press.

Weiner, H. (1992). *Perturbing the organism: the biology of stressful experience.* Chicago, IL: University of Chicago Press.

Psychosomatic Heart Disease: Role of Sympathetic and Sympathoadrenal Processes

G W Lambert
Baker Heart Research Institute, Melbourne, Victoria, Australia
T Dawood
Baker Heart Research Institute and Monash University, Alfred Hospital, Melbourne, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

Mechanisms Underlying Coronary Heart Disease and Myocardial Infarction and Possible Relationship to Stress and Behavior
Sympathoadrenal Function in Panic Disorder
Sympathoadrenal Function in Major Depressive Disorder

Glossary

Neurogenic essential hypertension	Elevated blood pressure initiated and sustained by a characteristic elevation in sympathetic nervous activity.
Major depressive disorder	A common and severe mental illness characterized by feelings of hopelessness, helplessness, and worthlessness; often accompanied by suicidal thoughts.
Panic disorder	A distressing and disabling condition with patients experiencing recurrent episodes of unexpected intense anxiety of sudden onset.

I was walking along the road with two of my friends. Then the sun set. Then the sky became a bloody red and I felt a tinge of melancholy, a sucking pain beneath my heart. I stopped, leaned against the railing, dead tired. Over the blue-black fiord and city hung blood and tongues of fire. My friends walked on and I stood again trembling with fright. And I felt as if a loud unending scream were piercing nature. – Edvard Munch

Mechanisms Underlying Coronary Heart Disease and Myocardial Infarction and Possible Relationship to Stress and Behavior

Over recent years, substantial advances have been made in delineating the causes of coronary heart disease. At a community level, the importance of high blood pressure, tobacco smoking, and abnormal blood lipids as causal factors is well established. Recent epidemiological studies also point to an association between psychosocial stressors and increased risk of development of acute myocardial infarction. Mechanistically, the relevant issues linking psychosocial stress and acute cardiac events are whether stress, in one of its numerous forms, can (1) lead to the development of hypertension, (2) cause or aggravate atherosclerosis, and/or (3) trigger a cardiac event in the presence of existing atherosclerosis.

Relation of Stress to the Development of Hypertension

The importance of hypertension in the development of left ventricular hypertrophy, myocardial infarction, heart failure, and sudden death is well established. Indeed, over 90% of patients with heart failure in the Framingham study had a history of hypertension. The longitudinal studies of Timio and colleagues, examining hypertension development and cardiovascular morbidity and mortality in cloistered nuns, provide persuasive evidence linking the stress of daily life to the development of high blood pressure. Essential hypertension is commonly neurogenic, initiated and sustained by the sympathetic nervous system. While activation of the sympathetic nervous system outflows to the kidneys, heart, and skeletal muscle in hypertensive patients has been well documented, a second neural mechanism, reduced reuptake of the sympathetic neurotransmitter, norepinephrine, is also operative. Reuptake of norepinephrine into sympathetic nerves after its release terminates the neural signal. A fault in transmitter inactivation augments the effects of sympathetic nerve traffic. For the sympathetic nerves of the heart, approximately 95% of released norepinephrine is recaptured, so that the heart is more sensitive than all other organs to impediments in transmitter reuptake. Other cardinal signs of stress common to patients with neurogenic essential hypertension include activation of brain noradrenergic pathways and epinephrine release from the heart.

Relation of Stress to the Development of Atherosclerosis

There exist numerous population studies indicating that psychological disturbances, particularly anxiety and chronic stress, can contribute to atherosclerosis development. In the workplace, insufficient self-regulation of workload, deadlines, and planning and direction of the work have been linked to risk of developing atherosclerosis. Low socioeconomic status or employment level is associated with increased plasma fibrinogen and von Willebrand factor concentrations. Endothelial function is considerably impaired following acute experimental stress.

Relation of Stress to the Triggering of a Cardiac Event

Cardiac sympathetic nerves are preferentially activated by mental stress It is important to realize that sympathetic nervous activity at rest and in response to stress is regionalized. For instance, with even relatively mild experimental mental stress, such as that induced in the laboratory by performing difficult mental arithmetic, the sympathetic nerves of the heart are markedly and preferentially activated, whereas activation of muscle vasoconstrictor sympathetic fibers predominates during the cold pressor test. Mental stress in patients with existing coronary artery disease has been demonstrated to cause anginal chest pain and inadequate blood supply to the heart. The effect of mental stress on cardiac sympathetic nervous activation is greater in the elderly. The importance of cardiac sympathetic activation in these contexts is highlighted by the demonstration, in experimental animals with narrowing of the coronary arteries, that direct electrical stimulation of the cardiac sympathetics is capable of causing electrical instability of the myocardium, triggering disturbances of heart rhythm and cardiac arrest.

General relation between activity of the sympathetic nerves of the heart and sudden death In a range of clinical contexts, stimulation of the cardiac sympathetic outflow has been demonstrated to contribute to myocardial infarction, ventricular arrhythmias, and sudden death. Activation of the sympathetic outflow to the heart is common in patients unexpectedly developing ventricular tachycardia and ventricular fibrillation. Furthermore, myocardial stunning due

to sudden emotional stress is associated with exaggerated sympathetic stimulation. Similarly, in patients with heart failure, there is a high level of stimulation of the cardiac sympathetic nerves, which is a proven determinant of sudden death. β-adrenergic blockers have been recently shown to increase survival in heart failure patients.

Direct relation between mental stress and sudden death in certain circumstances In a rare inherited heart condition, long QT interval syndrome, there is electrical instability of the heart muscle. Mental stress is one proven immediate cause of cardiac arrest in sufferers. Some research linking mental stress to sudden death is disputed because of disagreement over what constitutes a stress, and whether stress can be accurately measured. Recent research showing that rates of sudden, nontraumatic death in people with underlying coronary disease were markedly increased during catastrophic events is free of this criticism, as no finessing is needed in the psychological measurement of stress.

Sympathoadrenal Function in Panic Disorder

Debilitating episodes of recurring, often inexplicable anxiety afflict as many as 12% of the population. These attacks typically are unpleasant and may be accompanied by physical symptoms such as sweating, palpitations, tremor, and a sensation of suffocation. Given the prominent signs of sympathetic activation that occur during a panic attack, panic disorder patients find themselves at the nexus between cardiovascular and psychiatric medicine. It has until recently been felt that although panic disorder is distressing and disabling, it does not constitute a risk to life. Sufferers often fear that they have heart disease because of the nature of their symptoms, but have been reassured that this is not the case. Epidemiological studies, however, indicate that there is an increased risk of myocardial infarction and sudden death in patients with anxiety disorders. This increased risk extends even to premenopausal women, who in general have low coronary risk.

Somatic Sensitization as a Panicogenic Factor

Patients with panic disorder often exhibit hypervigilance to somatic sensations such as an irregular heartbeat or lightheadedness. Given the convergence of neural networks involved in processing viscerosensory and contextual information and their links with brain regions pivotal in the fear network, it is possible that somatic sensation may be capable of triggering panic attacks. Such a concept is not without precedent. It has been known for some time that the presence of abnormalities of heart rhythm, supraventricular tachycardia in particular, can pre-date and in fact cause panic disorder. More recently, a very high rate of development of panic disorder was observed in patients with implanted cardiac defibrillators. Indeed, factors at play here seem to include the presence of cardiac symptoms and the enhanced vigilance that accompanies them.

Sympathetic and Adrenomedullary Function in Panic Disorder

Resting whole body, cardiac, and muscle vasoconstrictor sympathetic activity is normal in patients with panic disorder. While adrenal medullary secretion of epinephrine, measured by isotope dilution, is typically normal in these patients, there occurs a demonstrable release of epinephrine from the heart. It is probable that during the surges of epinephrine secretion that accompany panic attacks the cardiac sympathetic neuronal vesicles become loaded with epinephrine that is continuously co-released with norepinephrine in the interim periods between attacks. *In situ* synthesis of epinephrine, through activation within the heart of phenylethanolamine-n-methyltransferase (PNMT), perhaps induced by repeated cortisol responses during panic attacks, may also occur. In experimental models of stress, PNMT activation in the heart and other extra-adrenal organs has been demonstrated.

Norepinephrine Reuptake in Panic Disorder

As mentioned previously, reuptake of norepinephrine into sympathetic nerves after its release terminates the sympathetic neural signal. A phenotypic defect in norepinephrine transporter function has been described in patients with panic disorder. Such an abnormality in neuronal norepinephrine reuptake could sensitize the heart to sympathetic activation. The mechanism underpinning the defect in norepinephrine transporter function, and whether treatment modifies the activity of the norepinephrine transporter, remains unknown.

Sympathetic and Adrenomedullary Function during a Panic Attack

During panic attacks triggered cardiac arrhythmias, recurrent emergency room attendances with anginal chest pain and ECG changes of ischemia, coronary artery spasm in attacks occurring during coronary

angiography, and myocardial infarction associated with coronary spasm and thrombosis have been documented. The recurrent episodes of unexpected intense anxiety that characterize a panic attack are accompanied by a pronounced activation of the sympathetic nervous system and adrenomedullary secretion of epinephrine. A concomitant elevation in blood pressure and heart rate also occurs. The characteristic chest pain that occurs during an attack is additionally accompanied by the release of neuropeptide Y into the coronary sinus. Neuropeptide Y is a potent vasoconstrictor with the capacity to cause coronary artery spasm.

Sympathoadrenal Function in Major Depressive Disorder

Central Nervous System Neuronal Activity

The etiology of major depressive disorder (MDD) has been linked, variously, to brain monoaminergic neuronal dysfunction, alterations in monoamine receptor sensitivity, and stress-induced activation of the hypothalamic-pituitary-adrenal axis. The lack of demonstrated relationships between different clinical presentations and biochemical abnormalities has hampered the development of sensitive biochemical markers for MDD. A major difficulty in commonly used peripheral indices of a neurotransmitter (or neurotrophic factor) as an indicator of brain dysfunction is that these compounds are also produced outside the brain. For this reason, venous blood is directly sampled from the brain to study central nervous system (CNS) neurotransmitter turnover, using central venous catheters placed high in an internal jugular vein. Under these circumstances, arteriovenous differences in neurotransmitter concentrations can be extrapolated to estimates of brain production. The observation of global (i.e., noradrenergic, dopaminergic and serotonergic) reduction in both subcortical monoamine turnover and cerebral glucose utilization in patients with MDD is consistent with a reduction in CNS neurotrophic support in patients with MDD.

Depression as a Risk Factor for the Development of Cardiac Disease

There is strong evidence that patients with MDD are at increased risk of developing coronary heart disease. This elevated risk is independent of classical risk factors such as smoking, obesity, hypercholesterolemia, diabetes, and hypertension. The association is present in males and females, across different age groups and in subjects living in different countries. The risk of coronary heart disease is increased 1.5- to 2-fold in those with minor/subsyndromic depression and 3- to 4.5-fold in subjects with major depression. To put this into perspective, the strength of the association between coronary heart disease and MDD is similar to that with smoking and hypercholesterolemia. Also conclusively demonstrated is the adverse effect of depression in patients with heart failure and in patients following myocardial infarction, which materially increases mortality.

Sympathoadrenal Activity in MDD

While the mechanism of increased cardiac risk attributable to MDD is at present uncertain, activation of the sympathetic nervous system may be of prime importance. Chronic stress has been linked with MDD. Examination of indices of sympathetic nervous function in MDD has yielded conflicting results. Some studies indicate a tendency for both the urinary excretion and cerebrospinal fluid level of norepinephrine and its metabolites to be diminished, whereas other reports document elevated plasma levels of norepinephrine and increased rates of norepinephrine spillover to plasma in patients with MDD. It is likely that sympathoadrenal activation is present only in a subset of patients with MDD. Indeed, Gold and colleagues recently demonstrated elevated norepinephrine levels in both the CSF and plasma of patients with melancholic depression.

In addition to possible effects on blood pressure and myocardial stability, sympathetically mediated neural vasoconstriction may exert metabolic effects in skeletal muscle, impairing glucose delivery to muscle and causing insulin resistance and hyperinsulinemia, and in the liver, retarding postprandial clearing of lipids and contributing to hyperlipidemia. A trophic effect of sympathetic activation on cardiovascular growth, contributing to the development of left ventricular hypertrophy, may also occur. Left ventricular hypertrophy is an independent risk factor for cardiovascular morbidity and mortality.

Further knowledge of the mechanisms responsible for generating cardiac risk may pave the way for novel and perhaps relatively simple therapeutic strategies (e.g., low-dose aspirin, β-blockers) to be administered to those with MDD in order to modify cardiac risk. Indeed, defining therapeutic interventions that reduce cardiac risk will be an important step forward in alleviating the burden of depressive illness on the community.

See Also the Following Articles

Antisocial Disorders; Depression and Coronary Heart Disease; Hypertension.

Further Reading

Bunker, S. J., Colquhoun, D. M., Esler, M. D., et al. (2003). "Stress" and coronary heart disease: psychosocial risk factors. *Medical Journal of Australia* **178**, 272–276.

Esler, M., Jennings, G., Lambert, G., et al. (1990). Overflow of catecholamine neurotransmitters to the circulation: source, fate, and functions. *Physiological Reviews* **70**, 963–985.

Gold, P. W., Wong, M. L., Goldstein, D. S., et al. (2005). Cardiac implications of increased arterial entry and reversible 24-h central and peripheral norepinephrine levels in melancholia. *Proceedings of the National Academy of Sciences USA* **102**, 8303–8308.

Lambert, G., Johansson, M., Agren, H., et al. (2000). Reduced brain norepinephrine and dopamine release in treatment-refractory depressive illness: evidence in support of the catecholamine hypothesis of mood disorders. *Archives of General Psychiatry* **57**, 787–793.

Leor, J., Poole, W. K. and Kloner, R. A. (1996). Sudden cardiac death triggered by an earthquake. *New England Journal of Medicine* **334**, 413–419.

Matsuo, T., Suzuki, S., Kodama, K., et al. (1998). Hemostatic activation and cardiac events after the 1995 Hanshin-Awaji earthquake. *International Journal of Hematology* **67**, 123–129.

Musselman, D. L., Evans, D. L. and Nemeroff, C. B. (1998). The relationship of depression to cardiovascular disease: epidemiology, biology, and treatment. *Archives of General Psychiatry* **55**, 580–592.

Rosengren, A., Hawken, S., Ounpuu, S., et al. (2004). Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study. *Lancet* **364**, 953–962.

Schlaich, M. P., Lambert, E., Kaye, D. M., et al. (2004). Sympathetic augmentation in hypertension: role of nerve firing, norepinephrine reuptake, and angiotensin neuromodulation. *Hypertension* **43**, 169–175.

Spieker, L. E., Hurlimann, D., Ruschitzka, F., et al. (2002). Mental stress induces prolonged endothelial dysfunction via endothelin-A receptors. *Circulation* **105**, 2817–2820.

Timio, M., Saronio, P., Venanzi, S., et al. (1999). Blood pressure in nuns in a secluded order: a 30-year follow-up. *Mineral and Electrolyte Metabolism* **25**, 73–79.

Wilkinson, D. J., Thompson, J. M., Lambert, G. W., et al. (1998). Sympathetic activity in patients with panic disorder at rest, under laboratory mental stress, and during panic attacks. *Archives of General Psychiatry* **55**, 511–520.

Relevant Website

www.beyondblue.org.au *beyondblue* is a national, independent, not-for-profit organisation working to address issues associated with depression, anxiety and related substance misuse disorders in Australia.

Psychosomatic Medicine

T Theorell
Karolinska Institutet, Stockholm, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by T Theorell, volume 3, pp 304–309, © 2000, Elsevier Inc.

Definitions
Link with Stress Physiology and Epidemiology
Social Environment
Gene Activation
Psychosomatic Attitude
Difference between Behavioral Medicine and Psychosomatic Medicine
Brain Anatomy and Physiology
Studies of Life Changes
Coping and Behavior Patterns
Psychosomatic Treatment
Liaison Psychiatry
Final Comments

Glossary

Alexithymia	Not reading emotions; an inability to identify and discriminate between specific emotions.
Amygdala	Paired almond-shaped nuclei that are part of the limbic system of the brain and are concerned with fear processing.
Anabolism	Rebuilding and repairing of tissues.
Cushing's syndrome	Hyperfunction of the adrenal cortex.
Gastric mucosa	The lining of the stomach wall.
Hippocampus	A paired seahorse-shaped structure that is part of the limbic system of the brain concerned with memory.
Hypothalamic-pituitary-adrenocortical (HPA) system	One of the two main stress response systems of the brain, which subserves the neural (neuroendocrine) control of glucocorticoid secretion by the adrenal glands. Among other actions, the glucocorticoids facilitate energy mobilization.

Definitions

When the concept of psychosomatic medicine was introduced in science it was founded in a medical discipline: psychiatry. Medical conditions (i.e., diagnoses) that were characterized by objective changes and that could be induced by psychological states were defined as psychosomatic. This early definition of psychosomatic included duodenal ulcer, ulcerative colitis, bronchial asthma, essential hypertension, rheumatoid arthritis, thyrotoxicosis, and neurodermatitis. Later definitions have been more focused on the individual's total state. According to such a definition, a psychosomatic condition is a state with somatic and psychological symptoms. The direction of causation is often difficult to establish. Thus, there may be psychological processes influencing somatic mechanisms as well as somatic phenomena that influence psychological states. The latter phenomenon has also been labeled somatopsychic.

Two early authors who were using the old way of defining psychosomatic were Franz Alexander (who was a psychoanalyst and a pupil of Sigmund Freud) and Flanders Dunbar. An important element in the efforts to diagnose conditions in the early days was the specificity concept. Specificity was used to label specific psychological processes that would be associated with specific psychosomatic diagnoses, e.g., a repressed cry in relation to an acute attack of bronchial asthma or an oral frustration during early childhood in relation to duodenal ulcer development in adult years.

A famous example of organic conditions that could be induced by psychological states was observed in the patient Tom, who had a gastric fistula. The fistula made the gastric mucosa available to the eyes of the researchers, which enabled the internists Harold Wolff and Stewart Wolf to conduct psychological experiments: they could study Tom's gastric mucosa while different states were induced in him. When long-lasting states of anger were induced, small ulcers arose, which were associated with elevated acid production, increased motility, and reddening of the mucosa. The opposite condition, a pale and passive mucosa with very little production of acid, was observed when Tom was in a state of long-lasting withdrawn and passive sadness.

Recently, discussion about peptic ulcer etiology has focused on *Helicobacter pylori*, a bacterium that is present in the gastric mucosa of most, but not all, patients with peptic ulceration. *Helicobacter pylori* seems to be necessary for the development of most peptic ulcers. Nevertheless, stress-induced variations in production of acid could increase the likelihood of ulcer development, especially in critical care situations. That is, the effects of emotional or stress-induced changes in gastric acid production could interact with *Helicobacter pylori* to cause gastric ulceration in some individuals. Robin Warren and Barry Marshall were awarded the 2005 Nobel Prize for the discovery of *Helicobacter pylori*.

In the more modern and wider definition of psychosomatic, the attitude toward the patient's illness is central. If any condition could be more or less psychosomatic, all caregivers should ask themselves whether there are aspects of their patient's condition that could be psychosomatic.

Link with Stress Physiology and Epidemiology

The influence of stress physiology on theoretical modeling in psychosomatic medicine became evident when internists became interested in the field. When psychiatrists collaborated with physiologists and internists, some of the basic mechanisms were revealed, explaining why allergic reactions, wide swings in diabetes mellitus, worsening of hypertension, and many other conditions could be partly psychosomatic.

The next step in the scientific development of the field was taken by epidemiologists, who made it possible to describe psychosomatic conditions in persons outside the doctor's office: the total illness panorama. In this way it became possible to identify those who suffer from psychosomatic conditions but do not consult a caregiver.

Studies of extreme stress were initially an important part of the psychosomatic field, and this is also where the link between stress and psychosomatic medicine was established. In Style's general adaptation syndrome (GAS), stress was seen as the general reaction to a nonspecific challenge or adverse condition that was labeled a stressor (a factor that induces stress). According to this definition, stress is a reaction, and it is a general one. It is evident that this reaction induces arousal, which means that energy is mobilized. There are several parallel phenomena, all of which aid the body in the physical fight-or-flight response. Examples include lowered excretion of water and salt, decreased sensitivity to pain, and decreased inflammatory response in infections. Because energy mobilization (resulting in an elevated blood concentration of glucose and free fatty acid) has the highest priority, anabolism is downregulated. Anabolism is central to the body's central defense of all the organ systems that need constant rebuilding and restoration. If this goes on for a long time (e.g., several months), increased sensitivity to physical and psychological stress in bodily organs is the ultimate result.

This represents a more general theory in contrast to the specificity theory mentioned earlier. Another consequence of long-lasting demands for energy mobilization is that endocrine systems may change their regulatory patterns. One consequence could be that the production of thyroid hormones is stimulated. Thyroid hormones stimulate a long-lasting upregulation of the metabolic rate.

Disturbed regulation of the hypothalamic-pituitary-adenocortical system (the HPA axis) has also been described in relation to long-lasting arousal. There seem to be at least three kinds of regulatory disturbances in the HPA axis:

1. Lack of inhibition of stimulation. Normally, cortisol inhibits higher centers in the HPA axis when it has been elevated for a long time. When this inhibition does not take place in the normal way, the levels tend to be elevated constantly.
2. Lack of response to stimulation of the HPA axis. This means that plasma cortisol levels are low and that the normal activation does not take place when the HPA axis is stimulated artificially.
3. Unstable levels. In this kind of disturbance, the plasma cortisol levels may be low at rest, but certain situations (that are specific to the individual) trigger a strong HPA axis reaction that may seem to be out of proportion to the importance of the trigger.

The first kind of regulatory disorder is typical of Cushing's syndrome and could be one alternative regulatory disturbance in psychiatric depression. Severe long-lasting psychiatric depression may be associated with constantly high cortisol levels; in most cases, however, the circadian rhythm is preserved.

The second kind of disorder has been shown to exist in a subgroup of patients diagnosed with chronic fatigue syndrome (CFS). Because at least one other group of patients, those with fibromyalgia, share some symptoms that are typical of this regulatory disorder with CFS patients, it is possible that part of the myalgia problem is that they also suffer from this disturbance. It should be pointed out, however, that CFS and fibromyalgia also have differences in symtomatology and that it is not known to what extent psychosomatic processes contribute to them.

The third kind of regulatory disorder is found in posttraumatic stress disorder (PTSD). The key element in the etiology of PTSD seems to be that the long-lasting extreme arousal level increases the number of cortisol receptors of the cell surfaces and also increases their sensitivity. Because the sensitivity is so high, only low concentrations of cortisol are needed to keep the system working at rest. This may be the reason why low serum cortisol levels are found at rest in patients with PTSD.

Disturbances in endocrine functions represent a more general theory that has been of central importance to psychosomatic medicine. Physiology has also contributed to the specificity theory, however. A recent contribution to cardiovascular psychosomatic medicine has been the hostility hypothesis, which states that subjects with a high degree of hostility and cynicism are more likely to develop a myocardial infarction at a relatively young age compared to other subjects. Hostility has been linked to disturbances in the production and blood concentration of serotonin.

Psychosocial processes are mostly not capable of inducing organic illnesses by themselves. There are other pathogenic processes involved that may have a genetic origin. One example that has been discussed recently is the role of the immune system in the etiology of thyrotoxicosis. Excessive production of some of the immune system's interleukines causes abnormal regulation of function of the thyroid gland, resulting in either thyrotoxicosis or hypothyroidism. Still, long-lasting life changes do have a role in facilitating the onset of thyrotoxicosis.

Social Environment

James P. Henry was one of the most active physiologists taking the position that the social environment has to be taken into account in the study of psychosomatic medicine. He represented a group of researchers who increased the understanding of psychosomatic medicine by taking into account both studies of colonies of animals whose social environments were manipulated and natural international variations across the human culture. One important finding was that blood pressure does not rise with increasing age in all cultures.

Gene Activation

Modern psychosomatic research has been influenced by the rapidly growing amount of knowledge about genes and gene activation. Genes that are of particular relevance to psychosomatic medicine are those that regulate monoamines such as thyroxin, noradrenalin, serotonin, and dopamine. This research explores mechanisms and importance of interactions between the environment and genes in the psychogenesis, for instance, of depression. An important landmark in these studies has been the prospective birth cohort study in New Zealand published by Caspi et al. A total of 1037 children were assessed every second year from age 3 up to age 15, every third year to age 21, and finally at the age of 26. Stressful life events occurring after the 21st and before the

26th birthday were assessed by means of a life history calendar. The risk of onset of depression after critical life events was shown to be related to a variant ("polymorphism") in the promoter (control) region of the serotonin transporter gene. One type of allele combination in this region is protective while another increases risk. These studies point at new directions in psychosomatic research.

Psychosomatic Attitude

The modern definition of psychosomatic is related to a treatment attitude toward the patients. The psychosomatic therapist is expected to consider social and psychological factors that may be of importance to the course of the patient's illness and to do this in all cases encountered. This of course means that any condition could be partly or totally psychosomatic. Such awareness should affect all medical treatment.

Difference between Behavioral Medicine and Psychosomatic Medicine

Behavioral medicine and psychosomatic medicine are similar to one another but are not identical. Behavioral medicine deals primarily with phenomena that are possible to record objectively, whereas psychosomatic medicine deals with intrapsychic phenomena as well. Psychosomatic medicine has been introduced mainly by psychiatrists and physiologists (biologists and medical doctors), whereas behavioral medicine has been introduced mainly by behavioral psychologists.

The therapeutic orientation has also been different in the two disciplines: whereas psychodynamically oriented treatments have dominated psychosomatic medicine, behavioral and cognitive psychotherapy has dominated behavioral medicine.

Brain Anatomy and Physiology

A major discovery in brain neurophysiology has changed the basic background of medicine. This is the fact that structural changes in the nerve cells of the hippocampus of the brain may occur as a consequence of traumatic experiences. Such discoveries have been made possible by new techniques for imaging the brain, building on principles such as magnetic resonance and positron emission. According to several studies, the volume of the hippocampus, a structure that is of great importance for the storage of basic memories, may diminish as a consequence of severe traumatic events that cause long-lasting changes in the concentration of several substances in the brain, with resulting PTSD.

Other neuroanatomical studies have shown that the interaction between emotions and cognition could be localized as traffic between specific structures such as the amygdala (representing nonspecific arousal and anxiety) and hippocampus (memory). Studies have shown that subliminal visual stimulation (pictures presented so rapidly that the person does not become aware of what they represent) may induce activity in the limbic system, e.g., the amygdala. It is known that there are direct connections between the visual cortex and the limbic system, and the cortex may therefore be activated at a later stage than the limbic emotional system. Similar connections exist for auditory stimuli. Such knowledge is of great importance in explaining the fact that subconscious processes may induce strong psychosomatic reactions.

Studies of Life Changes

The basic theory behind studies of life changes in psychosomatic research was Selye's nonspecific stress theory. It was later formulated by Holmes and Rahe as a theory of demands for adaptation to any change. A system of life change units was constructed on the basis of average scores from individuals in the normal population on demands for adaptation to specific life changes such as death of spouse, loss of job, and residential move. This line of research has shown that a group of subjects with an accumulation of important life changes during a short period of time will have an increased number of illness episodes during the subsequent period. This association holds on a group level but it is quantitatively relatively weak. It is when coping patterns, and also the perception of the life change by the subjects themselves, are taken into account that predictions get stronger. During later years, studies have shown that for certain types of illnesses and physiological reactions, a distinction could be made between positive and negative life changes – with positive changes in some cases inducing improvement and negative changes inducing deterioration in health. Studies of life changes have been important to the practice of psychosomatic medicine. Particularly in chronic illness rehabilitation, it is very useful to make systemic explorations of ongoing life changes. The patients immediately understand the value of this method, and the exploration may lead to increased insight, which could be very useful. Discussions regarding life changes that the patient has experienced with a concomitant recording of physiological functions may strengthen the insight further.

A sociological school of life change research, represented by Brown and others, has concentrated on the objective aspect of life changes. Special schemes for recording objective aspects of life changes are being

used. By analyzing the structural meaning (not the perception) of a life change, these authors have shown that there is an independent contribution of the change itself to the risk of illness.

Coping and Behavior Patterns

The way in which the individual perceives his or her life situation is of profound importance to his or her health. A basic concept in psychosomatic medicine is alexithymia, an inability to identify and discriminate between specific emotions. According to the theory introduced by Sifneos, primary alexithymia arises when emotions have not had the possibility to develop in the normal way during childhood, with resulting partial or total inability to differentiate specific feelings such as joy, anxiety, sadness, and rage. Secondary alexithymia, however, may arise during major life crises such as traumatic experiences or severe illness episodes. Patients with pronounced alexithymia may be very difficult to treat with verbal psychotherapy, because they do not have access to effective interpretations of what happens in their bodies in relation to emotionally arousing situations. Educational programs are presently being tested. A related concept is *pensée opératoire*, operative thinking, and a propensity to describe practical acts rather than emotions, e.g., in crisis situations.

Another aspect of coping that has been of particular importance to psychosomatic medicine is type A behavior. It was originally described as a behavior pattern that was elicited in certain situations. The cardiologists Friedman and Rosenman had seen it in younger male patients with coronary heart disease. In early stages of the scientific development it was discovered that there were several distinct components in this behavior, such as excessive drive and ambition (to manage difficulties better than others against the worst possible obstacles), hostility/feeling of rush, and excessive work orientation. Later researchers have criticized the definition of general type A behavior and stated that cynical hostility may be the key component. Evaluations of type A behavior intervention programs in secondary coronary prevention have shown that successful long-term programs (mostly lasting for at least a year) for reducing type A behavior may reduce the risk of developing new episodes of coronary heart disease.

Psychosomatic Treatment

Psychosomatic treatment follows different traditions in different countries. In several countries, e.g., Germany, there are several academically based clinics for psychosomatic treatment, mostly founded in psychoanalytical theory. In other countries, e.g., Sweden, it is believed that psychosomatic treatment should be part of good patient treatment on the whole, and therefore very few psychosomatic departments have been founded.

One of the key elements in psychosomatic medicine, from the time of Freud and his followers to today, has been alexithymia (see earlier discussion). Patients with long-lasting severe psychosomatic disturbances frequently suffer from early (childhood) disturbance of the emotional development. Chronic pain and chronic gastrointestinal functional disorders are examples of conditions that have been shown to be associated with traumatic childhood circumstances. It has been found that traditional verbal psychotherapy may be very difficult in such cases. Alternatives have been tried that could surpass verbal barriers. Such therapies could build either on direct bodily sensations or on indirect ways of helping the patients gain useful insight into emotion–body relationships. One such alternative is the use of arts that could be used either impressively or expressively. An example of the impressive use of music is guided imagery and music (GIM). GIM has been shown to have not only temporary but also long-term effects on endocrine systems. In an uncontrolled study of the use of dance, music, painting, or drama of patients with long-lasting (2 years in most cases) psychosomatic disorders, it was shown that the use of expressive arts can be helpful in gaining useful insight into the interplay between emotions and bodily symptoms. Experiences of art are often less guarded than verbal communications. In this study, the patients went through a catharsis phase of the treatment, mostly during the period 6 to 12 months after the start, in which strong emotions had been evoked and there was evidence of general physiological arousal. The art experiences were often helpful in creating turning points in the therapy.

Liaison Psychiatry

For a long time there has been a strong link between psychosomatic medicine and liaison psychiatry. In liaison psychiatry, specially trained psychiatrists work continuously in somatic medicine and have a special responsibility for patients with psychosomatic conditions. Liaison psychiatry includes areas in which psychological processes are of particular importance to healing of wounds and acceptance of transplants. The latter two areas are of particular importance to surgical wards. Studies during later years have shown that excessive long-lasting mobilization of energy prolongs the healing of wounds.

Final Comments

Psychosomatic medicine has developed rapidly since the introduction of the discipline. Specific theories along with general theories have been tested. In modern psychosomatic medicine there is growing awareness that the social environment is of great importance and also that the etiological processes are very complex. Interactions between physical, chemical, and biological agents (e.g., infections, such as *Helicobacter pylori*) and psychosocial processes cause long-lasting psychosomatic conditions.

See Also the Following Articles

Arthritis; Chronic Fatigue Syndrome; Hippocampus, Overview; Life Events and Health; Psychological Stressors, Overview; Type A Personality, Type B Personality; Ulceration, Gastric.

Further Reading

Brown, G. W. and Harris, T. (1978). *Social origins of depression: a study of psychiatric disorders in women*. London: Tavistock.

Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301**, 386–389.

Charney, D. S., Deutch, A. Y., Krystal, J. H., Southwick, S. M. and David, M. (1995). Psychobiological mechanism of posttraumatic stress disorder. *Archives of General Psychiatry* **50**, 394–430.

Dunbar, F. (1954). *Emotions and bodily changes: a survey of literature on psychosomatic interrelationships 1910–1953*. New York: Columbia University Press.

Friedman, M. and Rosenman, R. H. (1959). Association of specific overt behavior pattern with blood and cardiovascular findings. *Journal of the American Medical Association* **169**, 1286–1293.

Friedman, M., Thoresen, C. E., Gill, J. J., et al. (1986). Alteration of type A behavior and its effect on cardiac recurrences in post myocardial infarction patients: summary results of the recurrent coronary prevention project. *American Heart Journal* **112**, 653–655.

Henry, J. P. and Stephens, P. M. (1977). *Stress, health and the social environment: a sociobiological approach to medicine*. New York: Springer.

Holmes, T. H. and Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research* **11**, 213–218.

Jones, M. P. (2006). The role of psychosocial factors in peptic ulcer disease: beyond *Helicobacter pylori* and NSAIDs. *Journal of Psychosomatic Research* **60**, 407–412.

Kiecolt-Glaser, J. K., Page, G. G., Marucha, P. T., MacCallum, R. C. and Glaser, R. (1998). Psychological influences on surgical recovery: perspectives from psychoneuroimmunology. *American Psychology* **53**(11), 1209–1218.

Le Doux, J. (1999). *The emotional brain*. London: Phoenix.

Marty, P. and de M'Uzan, M. (1973). La pansée operatoire. *Revue Française de Psychoanalyse* **27**, 345–346.

Marucha, P. T., Kiecolt-Glaser, J. K. and Favagehi, M. (1998). Mucosal wound healing is impaired by examination stress. *Psychosomatic Medicine* **60**(3), 362–365.

McKinney, C. H., Antoni, M. H., Kumar, F. C. and McCabe, P. M. (1997). Effects of guided imagery and music (GIM) therapy on mood and cortisol in healthy adults. *Health and Psychology* **16**, 390–400.

Nemiah, J. C. (1996). Alexithymia: present, fast and future. *Psychosomatic Medicine* **58**, 217–218.

Rubin, R. T., Poland, R. E., Lesser, I. M., Winstone, R. A. and Blodgett, A. L. (1987). Neuroendocrine aspects of primary endogenous depression. I. Cortisol secretory dynamics in patients and matched controls. *Archives of General Psychiatry* **44**, 328–336.

Selye, H. (1950). *The physiology and pathology of exposure to stress*. Montreal, Canada: Acta Inc.

Sifneos, H. (1967). Clinical observations on some patients suffering from a variety of psychosomatic diseases. In: *Proceedings of the Seventh European Conference on Psychosomatic Research*. Basel: Karger.

Stollman, N. and Metz, D. C. (2005). Pathophysiology and prophylaxis of stress ulcer in intensive care unit patients. *Journal of Critical Care* **20**, 35–45.

Taylor, G. J. and Bagby, R. M. (2004). New trends in alexithymia research. *Psychotherapy and Psychosomatics* **73**, 68–77.

Theorell, T., Konarski, K., Westerlund, H., et al. (1998). Treatment of patients with chronic somatic symptoms by means of art psychotherapy: a process description. *Psychotherapy and Psychosomatics* **67**, 50–56.

Williams, R. B. (1996). Coronary-prone behaviors, hostility and cardiovascular health: implications for behavioral and pharmaceutical interventions. In: Orth-Gomér, K. & Schneiderman, N. (eds.) *Behavioral medicine approaches to cardiovascular disease prevention*. Mahwah, NJ: Erlbaum.

Psychotherapy

G C Smith
Monash University, Clayton, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G C Smith, volume 3, pp 310–315, © 2000, Elsevier Inc.

Introduction
Psychotherapy as an Interpersonal Process
Modification of Feelings, Cognitions, Attitudes, and Behavior
Troublesomeness of Phenomena
The Person Seeks Help
Help Is Sought from a Trained Professional
What Is the Difference between Psychotherapy and Education?
Stress Management
The Schools of Psychotherapy
Methods of Delivery of Psychotherapy
The Efficacy of Psychotherapy

Glossary

Counseling	A professional interpersonal process that overlaps with psychotherapy but that has as its focus specific problems or adjustments.
Psychoanalysis	A theory of mental development embracing the concepts of the unconscious as a determinant of behavior, the use of defenses against the emergence of unconscious tendencies, and the phenomenon of unconscious transference of relationships with past objects to present relationships.
Psychoanalytic practice	The application of a psychotherapeutic method characterized by emphasis on free association, transference, and interpretation, aimed primarily at producing greater self-awareness.
Psychotherapy	An interpersonal process designed to bring about modifications of feelings, cognitions, attitudes, and behavior that have proven to be troublesome to the person seeking help from a trained professional.
Stress disorder	A mental disorder as defined by the International Classification of Diseases and the Diagnostic and Statistical Manual of the American Psychiatric Association.

Introduction

Psychotherapy is an interpersonal process designed to bring about modifications of feelings, cognitions, attitudes, and behavior that have proven to be troublesome to the person seeking help from a trained professional. This time-honored definition, attributed to Strupp, is a condensation of the key issues of psychotherapy. A wider view of psychotherapy includes the concept of helping a person to become his or her best possible self. Psychotherapy is a major component of most stress management programs and treatments for stress-related disorders.

Psychotherapy as an Interpersonal Process

That one person can influence the way another feels, thinks, and behaves is a fundamental assumption in most cultures. Such influence can be either good or bad. The notion that language is involved is implicit. Psychotherapy is a professional refinement of this language-based interpersonal process. It usually consists of a verbal exchange, but may involve writing or dramatic performance as well.

Modification of Feelings, Cognitions, Attitudes, and Behavior

Psychotherapy aims at the modification of existing feelings, cognitions (thoughts), attitudes, and behavior. Psychotherapy avoids the issue of whether these phenomena are inherited or acquired, and it takes as given that they exist and are potentially modifiable. Nevertheless, there is an implied recognition that there has already been a process at work creating these phenomena, with or without the conscious awareness of the individual. Psychoanalysis and its derivative psychotherapies take account of the origin of these phenomena, but focus on their current manifestation in the transference, that is, on what occurs between the patient and therapist.

Troublesomeness of Phenomena

Disturbances of feelings, thoughts, and behavior are universal experiences. Psychotherapy deals with people who feel troubled by their experience of these phenomena, that is, people who are stressed and distressed. They may feel it directly, for example, in the case of depressed mood, or indirectly, for example,

when an upsurge of aggression leads to relationship problems or trouble with authorities. Stressed people are usually troubled by disturbance of their daily function. For the physically ill, stress may be troublesome because of its impact on their illness. The relationship between mental and physical states is clearest in the case of clinical depression. Having an episode of clinical depression at some stage in life at least doubles the risk of developing a major medical disorder such as coronary heart disease, stroke, or diabetes. Once these disorders develop, the presence of depression doubles the mortality rate and greatly increases morbidity and health-care costs, as do poor social support and poor social integration.

The Person Seeks Help

Psychotherapy as a form of interpersonal influence depends on and makes use of the fact that the troubled person seeks out help. In psychotherapies based on psychoanalytic theory, this action is central to the concept of transference. The person seeking help has expectations of the therapist, based on both the realistic knowledge of what the therapist has to offer and on unrealistic expectations transferred to the person of the therapist on the basis of relationship with other key figures in the person's life. This illustrates the fact that there is much going on in a psychotherapeutic relationship that is unsaid yet represented in language at some level and experienced as feelings and attitudes.

Help Is Sought from a Trained Professional

What differentiates psychotherapy from other forms of positive interpersonal influence is that it is offered within a formal set of parameters with an explicit basis for its conduct. That does not necessarily mean that it is valid, but it does mean that it is open to scrutiny. The trained professionals come from a variety of backgrounds, including medicine (particularly psychiatry), psychology, social work, nursing, and education. Specific training in psychotherapy is usually taken as a postgraduate course. There are institutes within the field that are devoted to training in and the development of psychotherapy. It is this postgraduate training that produces the skills and attitudes that differentiates professional psychotherapy and professional counseling from the informal counseling that many professionals provide in their everyday work.

Another implication of the involvement of a trained professional is that a code of ethics applies. Such codes are likely to include statements about respect for patients or clients, their values, their beliefs, their uniqueness, and their right to self-determination. In particular, psychotherapists are obliged to respect and take account of cultural factors and to avoid discrimination on the basis of such factors and beliefs. There is a consensus that the practice of professional psychotherapy requires ongoing supervision of some sort, in which there can be scrutiny of the application of these ethical principles. Supervision also addresses adherence to protocols. A special feature of supervision of psychoanalytic psychotherapy is that there is also scrutiny of the therapist's reaction to the patient – called countertransference.

What Is the Difference between Psychotherapy and Education?

All forms of professional psychotherapy contain an educational component, as does counseling, even though education is not their aim. Even psychoanalytic psychotherapy, which deemphasizes the didactic component, provides a model of thinking that the patient or client may find that they have adopted. In the sense that psychotherapy is explanatory, it is educational. The difference between psychotherapy and education lies in the fact that psychotherapy deals specifically with those matters that trouble the person requesting it and uses techniques specific to that issue. However, it must be acknowledged that much of modern teaching, with its individualized programs and greater sense of responsibility for the wider concerns of the student, is psychotherapeutic in nature. Many educational institutions are recognizing this by introducing mental health studies into their training. This does not mean that educators will become professional psychotherapists; that would require postgraduate training.

Stress Management

Stress management is the name given to an intervention offered increasingly and to a wide variety of target populations, including those at risk of stress and those who have experienced it. Often given in a group setting, it most frequently includes specific psychotherapy techniques but also imagery and relaxation techniques and an educational component.

The Schools of Psychotherapy

Psychoanalytic Psychotherapy

Professional psychotherapy has its historical roots in the development of psychoanalysis at the end of the nineteenth century. Psychoanalysis as a practice is a form of psychotherapy. Psychoanalytic

psychotherapy is a term used to describe all forms of psychotherapy that are based on the principles of psychoanalysis. The expressions dynamic and psychodynamic can be taken as a synonym for psychoanalytic. The distinction between psychoanalysis and psychoanalytic psychotherapy is indistinct, but it centers on issues of frequency, technique, and goal. Psychoanalysis is usually more intense in frequency and length. Psychoanalytic psychotherapy uses modified analytic techniques and may mix them with nonanalytic techniques. Psychoanalytic psychotherapy is often more problem-focused and may be explicitly supportive.

The issue of whether a form of psychotherapy is described as supportive is a complex one. All psychotherapies are supportive in the sense that they offer a containing framework in which a patient or client can deal with his or her stresses and lessen his or her distress. The concept of nondirectiveness is more pertinent. In psychoanalytic psychotherapy, interventions that contain suggestions or directions are usually avoided because they may interfere with the achievement of greater self-awareness, a major goal of psychoanalytic practice.

As the field of psychoanalysis developed various theoretical streams, so there developed different schools of practice of psychoanalysis and its derivative psychotherapies. Brief dynamic psychotherapy is associated with the names Malan, Sifneos, Strupp, Davanloo, Mann, and Horowitz. Klein, Winicott, Balint, and Guntrip pioneered object relationships theory and practice. Kohut developed self-theory. Lacan emphasized the linguistic basis of the structure of the unconscious; schools of psychoanalysis based on his teaching have wide support in Europe and South America. Jung, with his emphasis on universal symbols, spawned persisting schools of psychoanalysis.

Psychotherapies Related to Psychoanalysis

Other theorists such as Carl Rogers, Aaron Beck, and Harry Stack Sullivan focused on aspects of the subjective and cognitive experiences with which psychoanalysis dealt. They and others developed forms of psychotherapy based on their theories.

Experiential therapy Client- or person-centered therapy, existential therapy, Gestalt therapy, and emotional-expressive therapies are the most well-known examples of experiential therapy. They have in common the assumption that human nature is growth-oriented, and they focus on subjective experience. For instance, client-centered therapy is described by Rogers as having parallels in existentialism. It emphasizes the autonomy of clients (and thus avoids use of the term patient). It also emphasizes the attitudinal qualities of the therapist; warmth, empathy, and genuineness.

Cognitive therapy Cognitive therapy assumes that people can develop automatic erroneous thought patterns and beliefs of which they are unaware and that they consequently think and behave in ways that are irrational. Albert Ellis developed the concept of cognitive restructuring in his rational emotive therapy. Beck developed a similar technique, which is the basis for that used by therapists today, in which the patient or client is helped to identify these automatic thoughts and to challenge them, using what has been described as a Socratic dialog. There are both psychoanalytic and phenomenological underpinnings to the therapy. Ryle has developed cognitive analytic therapy, which emphasizes the psychoanalytic component. Analyses of transcripts of therapists using cognitive therapies show that they incorporate elements of other therapies as well. The constructs of behavioral therapy are often so interwoven with those of cognitive therapy that the term cognitive-behavioral therapy has come to be used. Cognitive-behavioral therapy is a major component of stress management programs and of interventions for acute and chronic traumatic stress disorder. It is widely used in the treatment of psychiatric disorders in general.

Interpersonal psychotherapy Interpersonal psychotherapy is an eclectic blend of traditional clinical approaches that has its roots in the work of Sullivan and was operationalized by Gerald Klerman and colleagues. It focuses on current interpersonal issues, particularly those related to grief, interpersonal role disputes, role transitions, and interpersonal deficits. It is, thus, particularly suited to those experiencing stress. It is a brief form of therapy, originally used for the treatment of depression but now applied to many other forms of distress. It, together with cognitive-behavioral therapy, was found to be as effective as antidepressant medication for the treatment of moderate depression in a large study conducted by the National Institute of Mental Health in the United States. Psychodynamic interpersonal therapy emphasizes the use of psychodynamic techniques. It has been shown to be both efficacious and effective in the physically ill; that is, it works in practice as well as in experiments.

Behavioral Therapy

Behavioral therapy has its roots in classical learning theory, with which the names of Pavlov and Skinner are associated. It is also influenced by Bandura's theory of socially determined learning or modeling.

Watson and Wolpe's early clinical work has now been refined and operationalized into a therapy that focuses on current determinants of behavior and draws on the principles of learning to develop individual treatment strategies. Behavioral therapists are aware of the fact that they cannot avoid completely the issue of the meaning of the behavior targeted or the meaning of the relationship between the therapist and the patient.

Methods of Delivery of Psychotherapy

Psychotherapy can be delivered at all ages, and most types of psychotherapy are applicable to all age groups. Melanie Klein and Anna Freud were pioneers of child psychotherapy. There is an emerging discipline of old age psychotherapy.

Individual Therapy

Individual or one-to-one psychotherapy is the commonest application of psychotherapy, but psychotherapy in small groups is an important form.

Couples Psychotherapy

Couples psychotherapy, for those in a relationship, is used by therapists of most schools of psychotherapy. Many therapists use an eclectic approach, combining elements of various techniques. Sexual issues are often an important focus.

Family Psychotherapy

Family psychotherapy, or family therapy as it is now known, has its roots in psychoanalytic theory. It began with the work of Bateson and colleagues in the 1950s on the role of the family in cases of schizophrenia. Salvador Minuchin and colleagues developed techniques based on a psychoanalytic construct of the family, which focused on ways in which one member could be carrying the anxiety for the whole family. More recently, family therapy has been strongly influenced by systems theory. The construct of family is changing rapidly, and given that each member of a family may come with different needs, family therapists have to adapt their technique to the situation.

Group Psychotherapy

Group psychotherapy can be said to have originated during World War II. Psychoanalytically orientated psychiatrists such as Foulkes and Bion used it in England to manage stress-related cases of neurosis more efficiently. The concepts and techniques that they developed are influential today. The Tavistock Institute in the United Kingdom has pioneered models of applied psychoanalysis for work groups trying to understand their social task. Such work forms part of stress management programs. Other schools of psychotherapy have developed group techniques. In the United States, Rogers extended his client-centered therapy to its application in encounter groups. Lewin developed methods for use in nonclinical situations. Behavioral therapy (e.g., for phobias) can be delivered in groups.

The Efficacy of Psychotherapy

Studies on the efficacy of psychotherapy for stress disorders as defined by international classifications are limited in number. However, the vast literature on the efficacy of psychotherapy in general is highly pertinent because it addresses the issue of stress by implication. Those diagnosed as having depression or anxiety, the most frequently addressed disorders in the research literature, are likely to have declared stress to be their major symptom, and this will have been acknowledged by their clinician when translating symptoms into a named problem or disorder. Furthermore, even when a primary diagnosis of a stress disorder is made, it is highly likely that a simultaneous diagnosis of depression, anxiety, or substance abuse will also be made. Personality disorder is another common comorbidity. The diagnoses of depression and anxiety imply distress, and the seeking of therapy certainly does so. These conditions, rather than the stress, may be the targets of psychotherapy. However, the exploration of perceived stressors and their effects forms an important part of all psychotherapies.

Overview

Health-care providers and consumers in general seek evidence of the efficacy of psychotherapy and justification of its cost. The empirical study of a therapy as complex as psychotherapy, applied as it is to such a wide range of problems and people, is fraught with methodological difficulties. Nevertheless, these difficulties have been addressed sufficiently well for Lambert and Ogles, in their review of efficacy and effectiveness studies, to conclude that many psychotherapies have been shown to have demonstrable effects on clients in clinical trials that are statistically significant and clinically meaningful and that speed up the natural healing process as well as providing coping strategies for future use. Although some clients may achieve meaningful results after brief psychotherapy (21 sessions), many need at least 50 sessions. Other conclusions were that the effects of therapy tend to be lasting, that psychotherapy patients show gains that surpass those resulting from placebo

controls or no treatment, that psychotherapy can be cheaper than medication, and that clients treated with psychotherapy show 25% less medical service utilization. However, they also conclude that average positive effects mask considerable variability in outcomes, that not all are helped, and that some are harmed by inept or inappropriate application.

Systematic and Meta-Analytic Studies

These conclusions were based on studies conducted over the past 80 years. Many of the earlier ones were naturalistic studies of therapy and follow-up. Some used qualitative research of individual cases, a technique now much refined and becoming increasingly important in the study of human behavior and experience in general. These techniques are particularly useful for examining the process of psychotherapy and the experience of patients or clients. Many of the recent studies are experimental; that is, a research plan is developed and clients are recruited for participation and follow-up, with monitoring of therapists as well. Systematic overviews of experimental studies group those that meet strict criteria for inclusion, such as the requirement of random allocation of clients to a treatment or control arm (randomized controlled studies, RCTs). Both qualitative and quantitative reanalyses may be applied. The statistical technique of meta-analysis is often used. It has methodological problems, but it is generally accepted that it provides a useful way of summarizing studies, albeit one that must be supplemented by more qualitative analyses. Early meta-analytical studies found that the average treated person is better off than 80% of the untreated sample (effect size of 0.85 standard deviation units). Lambert and Ogles conclude that studies since then, which include those on other large samples and a reworking of the original study and use of more sophisticated statistical techniques, have supported that finding in general but find lower but still clinically significant effect sizes (0.4–0.6) for both specific and general measures of outcome. The analyses on which these conclusions are based include studies on a wide range of therapies, including all of the mainstream ones, for specific disorders such as anxiety and depression and for general problems such as those of stress and personality. A wide range of severity of psychopathology is included. By way of perspective, effect sizes are similar to those found for the use of antidepressants in depression and pharmacological treatment of anxiety.

With respect to stress and its disorders, Ehlers and Clark conclude that systematic reviews and meta-analyses show that, contrary to the commonly held view, single-session individual debriefing does not reduce distress and is not effective in preventing posttrauma symptoms. However, a course of up to 16 psychotherapy sessions may be more effective than supportive counseling in ameliorating the severity of acute symptoms and reducing subsequent posttraumatic stress disorder (PTSD), a chronic condition. Bradley et al., in their systematic review of psychotherapy for PTSD, conclude that substantial effects are found for psychotherapeutic intervention, with the majority of individuals showing improvement that was clinically meaningful; this was sustained for at least 6 months.

Which Therapy for Which Patient?

The question of which therapy is best for which client is only slowly being resolved. Relationship factors (trust, warmth, acceptance, and human wisdom) are probably crucial, even in the more technical therapies that generally ignore them. In fact, it is becoming clear that therapists in clinical trials as well as in practice are eclectic in their adaptation of therapy to suit client profiles. Lambert and Ogles conclude that differences in outcome between various forms of therapy are not as pronounced as might have been expected and that it is premature to favor the so-called empirically supported therapies over others. (This term refers to a movement that advocates that only those therapies with proven experimental support should be taught and funded.) Courses of psychotherapy may be shown to be effective in clinical trials, but they are not as efficient in clinical practice, partly because they are too short. Emerging naturalistic studies show a dose–response relationship such that long-term therapy is required, at least to the point of complete resolution of symptoms.

As previously discussed, stress and related disorders constitute risk factors for the development of physical illness and its outcome. Some work has been done on specific psychotherapy techniques for patients with physical and psychological comorbidity. Studies mainly focus on psychoeducational interventions. Health education and stress management programs produce significant positive effects on blood pressure, cholesterol, body weight, smoking behavior, physical exercise, and eating habits. Psychotherapy for depression in diabetes is effective in relieving symptoms and improving control. Early trials of psychotherapy in patients with cancer showed an increase in both well-being and survival time. Improved physical treatments for cancer and the greater availability of psychosocial treatments may explain why later studies are finding no differences in survival time, despite improvement in well-being and quality of life.

What Are the Active Ingredients of Psychotherapy?

The work of Pennebaker and his colleagues has shown that the writing down of accounts of stressful experiences is accompanied by better outcome than is achieved by writing about nonemotional matters. This supports other evidence that suggests that affective experiencing – catharsis – may to be an important ingredient of psychotherapy. Cognitive mastery, making sense of what is happening, also seems to be important. A focus on strategies for behavioral control is the third major factor identified so far. But there is an enormous amount of research that needs to be done if the active ingredients of psychotherapy are to be dissected out and the question "Which therapy for which client by which therapist and how much of it?" is to be answered. Consistent with this realization, researchers are returning to a focus on the ingredients of psychotherapy: what the therapist and patient actually do, and what goes on between them. Systematic reviews and meta-analyses will help answer the questions, but the medical model on which they are based is insufficient to address this most human of experiences, that of someone's seeking help from a trained professional and entering an interpersonal process designed to bring about the modifications of feelings, cognitions, attitudes, and behavior that have proved troublesome to the person or, indeed, an individual's seeking greater self-awareness as a way of fulfilling him- or herself better in his or her personal and social commitments.

See Also the Following Articles

Cognitive Behavioral Therapy; Coping Skills; Crisis Intervention; Defensive Behaviors; Family Therapy; Group Therapy; Reenactment Techniques; Trauma Group Therapy.

Further Reading

Ablon, J. S. and Jones, E. E. (2002). Validity of controlled clinical trials of psychotherapy: findings from the NIMH Treatment of Depression Collaborative Research Program. *American Journal of Psychiatry* **159**, 775–783.

Bloch, S. (1996). *An introduction to the psychotherapies*. Oxford: Oxford University Press.

Bradley, R., Greene, J., Russ, E., et al. (2005). A multidimensional meta-analysis of psychotherapy for PTSD. *American Journal of Psychiatry* **162**, 214–227.

Duffy, M. (ed.) (1999). *Handbook of counseling and psychotherapy with older adults* New York: John Wiley.

Ehlers, A. and Clark, D. M. (2003). Early psychological interventions for adult survivors of trauma: a review. *Biological Psychiatry* **53**, 817–826.

Elkin, I., Shea, M. T., Watkins, J. T., et al. (1989). National Institute of Mental Health Treatment of Depression Collaborative Program: general effectiveness of treatments. *Archives of General Psychiatry* **46**, 971–982.

Hollon, S. D., Thase, M. E. and Markowitz, J. C. (2002). Treatment and prevention of depression. *Psychological Science in the Public Interest* **3**, 39–77.

Kazdin, A. E. (2000). *Psychotherapy for children and adolescents: directions for research and practice*. New York: Oxford University Press.

Lambert, M. J. and Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In: Lambert, M. J. (ed.) *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th edn.), pp. 139–193. New York: John Wiley.

Markowitz, J. C. and Weissman, M. M. (2004). Interpersonal psychotherapy: principles and application. *World Psychiatry* **3**, 136–139.

Ong, L., Linden, W. and Young, S. (2004). Stress management. What is it? *Journal of Psychosomatic Research* **56**, 133–137.

Reinecke, M. A. and Clarke, D. A. (eds.) (2004). *Cognitive therapy across the lifespan.* Cambridge, UK: Cambridge University Press.

Roth, A. and Fonagy, P. (2005). *What works for whom? A critical review of psychotherapy research* (2nd edn.). New York: The Guilford Press.

Sandell, R., Blomberg, J. and Lazar, A. (2002). Time matters: on temporal interactions in long-term psychotherapies. *Psychotherapy Research* **12**, 39–58.

Smith, G. C. (2003). The future of consultation-liaison psychiatry. *Australian and New Zealand Journal of Psychiatry* **37**, 150–159.

Spiegel, D. (2002). Effects of psychotherapy on cancer survival. *Nature Reviews Cancer* **2002**(May), 383–389.

Westen, D., Novotny, C. M. and Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin* **130**, 631–663.

Psychotic Disorders

J Ventura
University of California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Ventura and R P Liberman, volume 3, pp 316–326, © 2000, Elsevier Inc.

Stress in Psychotic Disorders: Conceptual Framework
Stressful Life Events and Psychosis
Natural Disasters and Psychotic Disorders
Expressed Emotion as a Stressor
Treatment as Protection against Stress and Psychosis

Glossary

Brief psychiatric rating scale (BPRS)	One of the most widely used rating scales, nationally and internationally, for measuring psychiatric symptoms at baseline and for assessment of change over time. The BPRS allows for the rapid assessment of symptoms such as depression, anxiety, delusions, and hallucinations.
Cognitive appraisal	An individual's evaluative thought process that instills a life event or situation with a particular meaning.
Coping response	Cognitive and behavioral efforts used to master, tolerate, and reduce the demands of stressful events that tax or exceed an individual's resources.
Expressed emotion (EE)	Attitudes and feelings expressed by a relative during family interactions; an individual with a psychotic disorder would describe this as critical comments, overt or subtle hostility, or emotional overinvolvement in the individual's affairs.
Independent life event	A life event or situation that is not related to an individual's psychotic symptoms and is beyond his or her personal influence, for example, the death of a relative, the loss of a job because the entire company went out of business, or becoming the victim of an unprovoked attack.
Mediator	A variable that explains how, why, or when to expect a relationship between a predictor (e.g., life event) and the criterion (e.g., psychosis).
Prospective studies	Research in which the data are collected on an ongoing regular basis about all study participants, whether they are considered at risk for an illness or not and before the event being studied occurs (e.g., psychotic relapse).
Protective factor	Individual characteristics, which may be biological, intrapersonal, interpersonal, or learned, that reduce or eliminate the risk of developing a psychosis.
Stress	An adverse biological, physiological, or psychological state that develops when the demands on an individual exceed his or her capacity for adaptation.
Vulnerability to psychosis	The enduring characteristics of individuals, which might be genetic, social, behavioral, biochemical, cognitive, or related to brain structure, that increase their susceptibility to developing psychosis.

Stress is a concept on whose definition few scholars or researchers can agree. Yet there is little dispute that stressful events can influence the onset and course of major psychotic disorders. A 1977 vulnerability–stress model by Zubin and Spring proposes that a relationship exists among the level of inherited vulnerability, stress that exceeds a certain threshold, and the onset or relapse of psychotic illness. According to this vulnerability model, individuals with high biological vulnerability are susceptible to the exacerbation of psychotic symptoms even when they experience low levels of stress; however, if the biological vulnerability is low, then high levels of stress are required before symptoms emerge. Differences in the rates of psychotic disorders found in males versus females, urban versus rural areas, and high-stress versus low-stress households support the importance of environmental factors in understanding gene–environment interactions.

Several empirical tests of the vulnerability–stress model have shown that an increase in the frequency of life events occurs in the month before a psychotic exacerbation. However, not all studies have found a reliable association between stressful life events and psychosis. Some researchers believe that the relationship in schizophrenia is weak when compared to more striking findings in patients with mood disorders. Although there is a convergence of evidence that stressful life events are triggers of psychosis, stressful events alone are neither necessary nor sufficient to account for illness onset, nor are they necessary as precursors of relapse. Antipsychotic medication may provide protection against stress and raise the threshold of vulnerability. An interactive vulnerability–stress–protective factors model (Figure 1) hypothesizes the existence of mediators of stress, some of which are considered protective such as

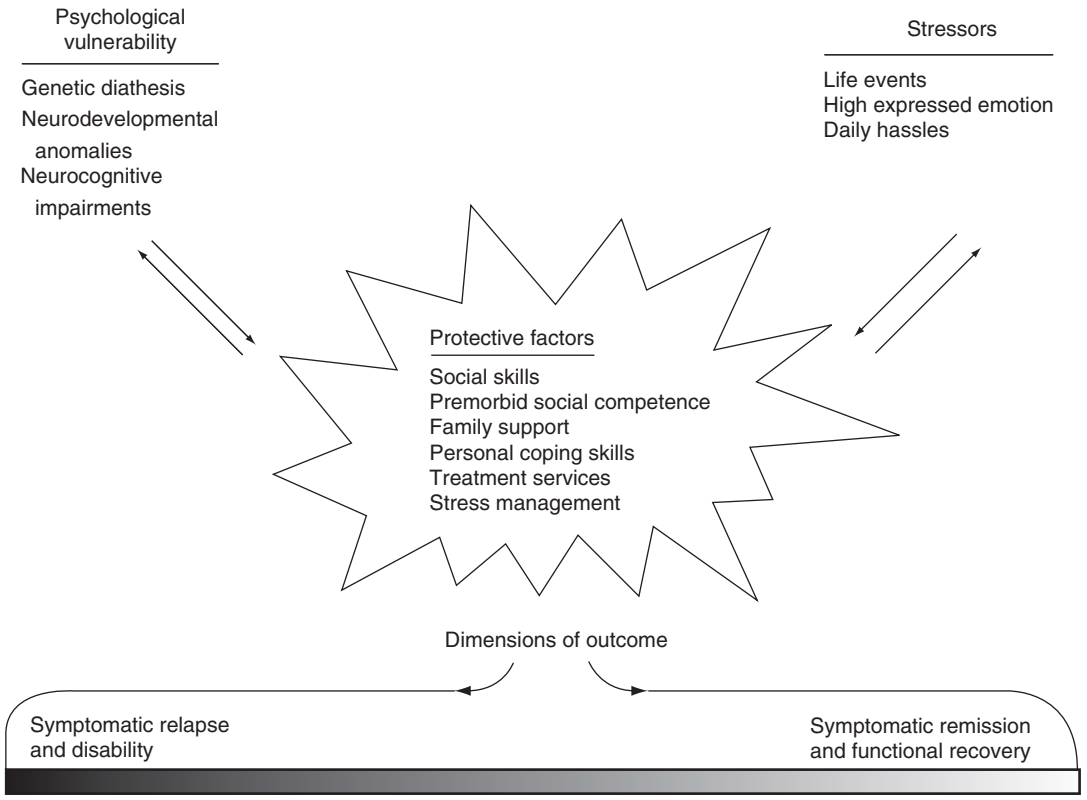


Figure 1 An interactive vulnerability–stress–protective factors model of risk factors for the onset and course of psychotic disorders.

cognitive appraisal of life events, social competence, social and familial support, and premorbid intelligence.

Stress in Psychotic Disorders: Conceptual Framework

All biomedical disorders, including those that manifest themselves by psychotic symptoms, are stress-related biological illnesses. It doesn't matter whether it is the heart or the brain that is involved as the organ from which the illness arises, stressors impinge on the organ's genetic vulnerability to trigger episodes of symptoms, exacerbations, relapses, and disability. Moreover, protective factors in the individual or in the individual's environment can mitigate or buffer the noxious effects of stressors on vulnerable people. When protective factors – such as social competence, social and family support, intelligence, availability, and accessibility to preventive interventions and good quality treatment – outweigh stressors and vulnerability, health may be sustained and disability forestalled.

In this article, we identify some of the stressors that have been implicated in the induction of episodes of illness, relapse, and poor outcome in psychotic disorders. However, it is important to place the role of stressors within the framework of the vulnerability–stress–protective factors model so as to understand why an individual may experience an episode of illness when exposed to a stressor at one time but not when exposed to the same stressor at another time. In addition, the vulnerability–stress protective–factors framework permits clinical scientists to explain the enormous interindividual variability in susceptibility to stressors – in short, why some people, some of the time, do not experience stress and exacerbations of psychosis in response to certain stressors, but others do.

Stressful Life Events and Psychosis

The life charts of Adolf Meyer, published in 1951, were among the first records of clinical observations that emphasized the importance of psychosocial stressors and the development of pathological conditions. Over the past 3 decades, many researchers have attempted to empirically document the nature of the relationship between stressful life events and psychotic illness. Such studies are based on the notion that higher frequencies of events translate into higher levels of stress for patients. These studies can be grouped into those that compared schizophrenia patients and other psychiatric groups, those that

compared schizophrenia patients with community controls, and those that examined the impact of stressful events on the onset and relapse of psychotic symptoms.

Most research has shown that, compared to schizophrenia patients, other psychiatric groups (usually depressed patients) reported either similar or significantly higher rates of stressful events prior to an episode. These differences were unexpected because there was an assumption that the rate of stressful life events in schizophrenia patients was higher than for patients with other psychiatric disorders. Schizophrenia patients are considered sensitive to even subtle changes in the environment. However, given that schizophrenia patients sometimes withdraw when stressed, their rate of life events may actually be lower than the rate for patients with other psychiatric disorders. Available evidence supports the conclusion that schizophrenia patients report fewer life events than other diagnostic groups. Furthermore, most studies have found that the strength of the relationship between stress and symptoms is more robust in patients with major depression than in patients with schizophrenia.

Most studies that used community controls as a comparison found that schizophrenia patients experienced a higher frequency of stressful events. However, some studies showed no difference. Interestingly, one prospective study found that normal controls actually had higher levels of stressful events than did schizophrenia patients. Several studies have shown an increased frequency of triggering life events in the month prior to the onset or relapse of schizophrenic symptoms. Even studies that did not find a sharp increase in the number of life events in the month prior to onset or relapse found that stressful life events had increased the risk of psychosis. There has been some investigation into whether life events play a more prominent role in the onset of psychosis than in relapse. Although there is no consistent evidence favoring a relationship between stress and onset (compared with stress and relapse), there is some indication that life events play a larger role in psychotic episodes that have a discrete (vs. insidious) onset.

Perhaps more relevant than the absolute frequency of events in patients or controls is the nature of the relationship between life events and psychotic illness. However, most of the studies of stress and psychosis focused on the frequency of events rather than the amount of threat posed. In fact, methods of measuring life event frequency (through open-ended interviews or life event inventories) are readily available. But, in some cases, events that were hoped for or expected (e.g., finding a suitable mate) but did not occur may also be perceived as stressful. Most researchers believe that the impact of a stressor is determined by an individual's cognitive appraisal and that the appraisal is altered in schizophrenia patients in a way that increases vulnerability to stress and psychosis; yet this important topic has received limited study. Perhaps the popular belief that most schizophrenia patients cannot accurately report their appraisal of life events dampens researchers' enthusiasm for comprehensive study.

Retrospective Studies of Life Events and Psychosis

In a classic and frequently cited 1968 British study, Brown and Birley interviewed schizophrenia patients and their families soon after a hospitalization regarding the 12-week period prior to the onset or relapse of psychosis. Schizophrenic patients reported an increased frequency of independent life events (independent of the insidious onset or occurrence of psychotic symptoms) in the 3-week period prior to the datable onset of their episode. Specifically, 46% of patients reported an independent life event in that 3-week period, compared with 14% for a sample of community controls during a comparable period. Except for the 3-week period prior to an episode, the life event rates between the two groups did not differ significantly. From these results, Brown and Birley concluded that life events play a role in triggering the onset or relapse of psychotic symptoms in schizophrenia.

The most ambitious attempt at a replication of the Brown and Birley findings was a 1987 multinational World Health Organization (WHO) study published by Richard Day and colleagues. The design of this WHO study was virtually identical to Brown and Birley's original work, including the use of retrospective reporting, although normal controls were not included. The life events data were collected in nine centers, five located in developed countries and four in developing countries. Day and colleagues replicated the finding that independent life events occur with increased frequency in the 3 weeks prior to psychotic episodes in schizophrenic patients. Although the lack of a comparison group is a serious concern, the consistency of these findings from different countries lends considerable support to the hypothesis that stressful life events may trigger episodes of schizophrenia.

Questions Concerning Retrospective Studies

Most of the classic research supporting the idea that stress triggers psychosis has been subject to criticism because the data was collected using retrospective methods, in which data regarding life events were

collected after onset or relapse of psychosis. Because retrospective data gathering relies on the patient's memory of life events and the date of onset of psychotic symptoms, it is subject to considerable report distortion, particularly for events of smaller magnitude. Such distortion is likely to bias results when patients (or their families) are searching for an explanation for a known outcome, such as a psychotic episode. In addition, there is the issue of bias in the dating of the onset or relapse of psychosis in these studies. Such a bias favors finding a relationship between stress and psychosis even if there was none.

The prospective studies of the mid-1980s and 1990s, which used a follow-through design in the regular collection of life events data, were aimed at correcting the methodological flaws of the prior retrospective studies. In some studies, life events data were collected as often as biweekly, thereby minimizing errors associated with memory for stressors that preceded relapse. In addition, each patient's symptoms were assessed frequently, minimizing the distortion of an association between the presence of symptoms and certain events.

Prospective Studies of Life Events and Relapse

The first well-designed prospective study of schizophrenia outpatients addressed the question of whether a significant increase in symptoms occurs after independent life events rather than evaluating the frequency of life events in the month preceding symptom relapse. This study, which collected life events data on an ongoing regular basis, focused on the average effects of life events on illness. An increase in negative symptoms was found in response to stressful events but not an increase in positive psychotic symptoms. Based on this stringent test of the relationship between stress and psychosis, Jean Hardesty and colleagues concluded that independent life events have minimal bearing on florid psychotic symptoms in schizophrenia.

A 1989 prospective study by Ventura and colleagues introduced additional methodological improvements in data collection, including (1) careful monitoring of the amount of antipsychotic medication received, (2) the use of both an open-ended style of interviewing and a structured life events inventory, and (3) the classification of life events as independent only when they were both independent of the patient's illness and not within the patient's ability to influence them. Ventura et al. found, as had Brown and Birley, and Day and colleagues, an increase in the number of independent life events in the 4-week period before a relapse or significant exacerbation of positive symptoms; 5 of 11 (45%) patients had at least one independent life event in the month before relapsing, compared with 1 of 11 (9%) during a comparable period without relapse. Thus, a well-designed prospective study supported the conclusion that stressful life events play a role in triggering psychotic symptoms.

Stress Sensitivity

Researchers and clinicians alike have theorized that schizophrenia patients are sensitive to even very small negative changes in the environment. A series of studies using the Experience Sampling Method (ESM) has provided some evidence for the validity of this hypothesis. The ESM requires that patients wear a beeping watch that alerts them 10 times a day to record real-life activity and their emotional response over a 6-day period. Using the ESM, researchers in the Netherlands found that schizophrenia patients, compared to high-risk relatives and controls, are more likely to experience negative affect in response to minor life events. This negative affectivity is considered a vulnerability marker for future psychosis. In addition, using the ESM the researchers recorded life events in sets of twins, in which one twin scored high on a measure of psychosis proneness and the other scored low. The advantage of this method of research is that the well twin acts as a genetic control for the psychosis-prone twin. Twins who were prone to experiencing psychotic-like experiences were more likely to show sensitivity to minor stressful life events. These findings further support the hypothesis that stress sensitivity acts as a risk factor for psychosis.

General Conclusions on Life Events and Psychosis

Interestingly, similar conclusions about the relationship between stress and psychosis can be drawn from sophisticated prospective studies and the retrospective studies. There is strong evidence from both retrospective and prospective studies to support the triggering hypothesis, even though some studies were able to confirm only an increased risk of psychosis from stressful events. Although life events influence psychotic episodes, the occurrence of stressful life events is usually not sufficient to precipitate an increase in psychotic symptoms. Furthermore, life events are not always necessary for relapse to occur. Some studies found that more than one-half of relapsing patients did not experience a prior life event. Other factors, such as medication reduction/discontinuance or biological factors, may also contribute to the risk of relapse. These conclusions attest to the potential importance of mediators such as, cognitive appraisal and coping efforts, which produce

differences in individual response to stress within groups of psychotic patients. The finding of increased negative symptoms in response to stressful events, instead of psychosis, indicates that not all patients respond to stress in the same way. Prospective research has also shown that life events do not have to be large in magnitude; both major and minor life events are predictors of symptom exacerbation and relapse in schizophrenia. Furthermore, stress sensitivity to daily hassles and minor life events has been associated with increased subjective distress and potential risk for future psychosis in schizophrenia.

Cognitive Appraisal of Stressful Events

Understanding the link between stressors and subsequent psychotic symptoms in schizophrenia may entail looking more closely at potential mediators such as patients' cognitive appraisal of life events. According to stress researchers, an individual's appraisal of an event always should be measured because appraisal determines the stressful nature of the event. In one stress–illness model, the perception of the stressor is schematically represented as a polarizing filter that can magnify or diminish the stressfulness of an event. Richard Lazarus and Susan Folkman hypothesized that appraisal is the first element of a two-phase cognitive-phenomenological process of coping and thereby is a factor in determining how much stress an individual experiences. After decades of research, Lazarus continued to assert that an individual's appraisal is necessary and sufficient in the generation of emotions in response to stressful events. Appraisal thus is relevant to understanding the relationship between stressful life events and subsequent illness.

Research suggests that schizophrenic patients, in particular, might respond in idiosyncratic ways to life events. Appraisal may play a larger mediating role in the relationship between life events and psychotic symptoms in schizophrenia patients than it does in other populations. Both psychoanalytic and cognitive information-processing theories suggest that schizophrenia patients rely on narrow and idiosyncratic meanings of events. Heightened sensitivity to environmental stimuli, coupled with cognitive distortions and psychotic thinking (e.g., persecutory ideas or referential thoughts to either perceived or existing threats) could result in increased stress. Research has shown that schizophrenia patients may overestimate the threatening nature of everyday events or situations and underestimate the effectiveness of their problem-solving ability. This tendency for altered appraisal may increase in patients as their condition worsens.

Can Stress Cause Psychotic Illness?

Although some investigators have argued that stressful life events can cause psychotic episodes in schizophrenia patients, others have argued that life events can only act as triggers of episodes. Central to this issue is whether psychosis can occur in individuals without inherited vulnerability and from stressful events that are definitely independent of the illness. Exposure to wartime combat, natural disasters, and other severe traumatic events that are clearly independent of the individual's influence have been shown to increase the risk of developing psychosis. According to one theory, psychosis can occur even in individuals without inherited vulnerability characteristics, especially if the stressor is severe. However, many individuals have been exposed to all types of stressful situations and catastrophic events and only a handful become psychotic. Some theorists believe that individuals who became psychotic must have some type of underlying vulnerability to illness and that therefore the stress merely acted as a trigger. The possibility also exists that the role of life events may be indirect and occur in conjunction with other risk factors.

Stressful Life Events and Depression in Psychosis

The presence of clinical depression in schizophrenia patients has generated a great deal of interest because of its significance as an early indicator of an impending psychotic relapse, association with an increased risk for suicide attempts, and relationship to psychiatric rehospitalization. Several theories have been put forward to explain the occurrence of depression in schizophrenia. Depression has been considered part of the postpsychotic phase when it occurs after a psychosis remits and as a reaction to the realization that the individual is in trouble or that his or her life has fallen far short of prior expectations. Other researchers view depressive symptoms as part of the acute psychotic process because it has been found to occur simultaneously with the exacerbations of positive symptoms. Depression may be a part of the patient's biological diathesis to schizophrenia and therefore an integral part of the disorder or an aspect of familial genetic liability to disturbances in mood. Some researchers view depression as an unwanted side effect of antipsychotic medication. However, observations by Eugen Bleuler and Emil Kraepelin of depression in schizophrenia during the preneuroleptic era suggest that depression is not simply secondary to antipsychotic medications. In any event, according to the vulnerability–stress–protective factors model, psychosocial stressors such as life events may be expected to increase the chances

of an onset or exacerbation of depressive symptoms in psychotic patients.

Schizophrenia patients have reported a higher frequency of stressful life events during the 6 months before a depressive exacerbation and during the 12-month period prior to a postpsychotic depression. Early parental loss (before age 17) was also found to be more common in depressed than nondepressed chronic paranoid schizophrenia patients. In chronic schizophrenia patients, increased levels of depression have been found to be associated with psychosocial stressors, such as unemployment. Although much of this previous research had some methodological weaknesses, recent work with sound methods confirmed the earlier findings. A well-designed 2000 prospective study using survival analysis showed that stressful life events act as triggers of depressive exacerbation in recent-onset schizophrenia patients. These studies suggest that individuals with schizophrenia who experience undesirable life events may have a vulnerability to depression similar to mood disorder patients and individuals without psychiatric disorders. Thus, stressful events may play an even broader role in psychosis than is typically theorized by the vulnerability–stress–protective factors model.

Natural Disasters and Psychotic Disorders

There are various types of stressful life events that can have a negative impact on schizophrenia patients, ranging from minor daily hassles to major life events. These include catastrophic events such as natural disasters. One reliable predictor of psychiatric morbidity is the degree of exposure to the devastating physical consequences that follow natural disasters. Research on natural disasters showed consistently that the existence of a prior psychiatric disorder (or symptoms) is a risk factor for postdisaster psychopathology. Preexisting disorders such as major depression, posttraumatic stress disorder (PTSD), anxiety disorders, and alcohol dependence have been associated with postdisaster morbidity. One study found 98% of individuals with a prior psychiatric disorder had a diagnosable postdisaster psychiatric disorder, compared with 25% of postdisaster diagnosis in the group without a previous diagnosis. Most studies have been retrospective and therefore subject to recall bias, but these results have been supported in prospective research.

Contrary to what is predicted by the vulnerability–stress–protective factors model, most anecdotal accounts and a few studies indicated that psychiatric patients remain clinically stable or even improve during or after natural or human-made disasters (e.g., hurricanes, ice storms, nuclear power plant accidents, and missile attacks). During Hurricane Iniki, psychiatric inpatients in Hawaii responded with increased independence, defined as decreased requests for therapy. Similarly, during SCUD missile attacks on Israel throughout the Persian Gulf War, many psychiatric inpatients were described as having coped well, quietly entering sealed concrete bunkers and easily donning gas masks. During a state hospital fire, the patients assisted the hospital staff in the efforts to control the blaze.

In contrast to these prior reports, patients with psychotic disorders in some studies of disasters demonstrated evidence of heightened stress reactivity. On January 17, 1994, a 6.8-magnitude earthquake struck Northridge, California, and provided our research group with an unprecedented opportunity to study the impact of a natural disaster on psychotic patients. Our sample consisted of patients who were previously diagnosed as having a psychotic disorder and were active participants in research projects. As hypothesized, the degree of exposure to the earthquake was a statistically significant predictor of the amount of distress experienced. Schizophrenia patients reported higher levels of avoidance symptoms than did controls when assessed postearthquake. However, the groups did not significantly differ in their mean levels of intrusion symptoms. The specific elevation of avoidance symptoms is noteworthy because the early appearance of avoidance and emotional numbing has been found to be the strongest predictors of disaster-induced psychiatric morbidity, including PTSD. BPRS assessments were compared for patients for whom 1-week preearthquake and 1-week postearthquake ratings were available. We found a significant increase in postearthquake dysphoria but not in psychotic symptoms. In response to the September 11, 2001, terrorist attacks, there was no difference between inpatients in New York who did and inpatients who did not have the opportunity to directly view the disaster through downtown-facing windows. However, patients with a schizophrenia spectrum diagnosis showed evidence of worsening symptoms compared to those with affective disorder or other diagnoses. Although there is some evidence of confirmation of the stress reactivity hypothesis in response to disasters, the stress reactions of psychotic patients are not always consistent and might depend on the amount of exposure, type of disaster, and the setting.

Expressed Emotion as a Stressor

High expressed emotion (EE), the attitudes and feelings expressed by a family member to a patient with a

psychotic disorder, is the most potent stressor precipitating relapses. EE is a misnomer because it does not refer to emotional expressiveness but, rather, to very specific and operationalized types of attitudes and feelings that are associated with stressful interactions within a family living with a severely mentally ill person. These attitudes and feelings take the form of critical comments, hostility, and emotional overinvolvement. Critical comments are complaints about specific behaviors or symptoms of the mentally ill family member, whereas hostility is generally directed at broader and more global personality characteristics. Emotional overinvolvement is overprotectiveness, intrusiveness, self-sacrificing behavior, and hovering over the patient by the family member. Conversely, low EE, which is associated with lower than average relapse rates, may subserve a protective function by insulating the individual from stressors with positive comments, positive reinforcement, and warmth.

Family members' levels of EE are reliably associated with patients' risks of relapse in the subsequent 9–12 months after a psychotic episode. Patients living with or having frequent contact with one or more high EE relatives have relapse rates 2–3 times greater than patients living with or having contact only with low EE relatives. This increased risk is present despite adequate treatment with antipsychotic medication. In a recent meta-analysis of 27 studies, the average rate of psychotic relapse in patients living with one or more high EE relatives was 52%, compared to 22% in patients coming from low EE families. The probability that this result was due to chance is virtually zero. This association of EE and relapse has been found for relapses in bipolar disorder, schizophrenia, and major depression.

Although there is no evidence whatsoever that EE or families have an etiological role in the first episode or development of psychosis, there is solid evidence that there is a causal link between EE and relapse. More than 15 treatment studies in several countries have shown that behavioral, social learning, and psychoeducational interventions that are well-organized, structured, and focused on teaching family members practical methods for understanding and coping with psychosis lead to reductions in EE and in relapses. Thus, the treatment research provides experimental evidence for the role of EE in stress-related relapse.

The association between EE and relapse in psychosis has been consistently found in developing countries (e.g., India) and developed countries (e.g., United States, United Kingdom, and Italy); however, cultural factors do play a role in the proportion of families from various cultures who are high versus low EE. For example, the proportion of high EE families is significantly lower in rural India than in urban India and is lower among unacculturated Mexican American families than among Anglo-American families. It is likely that cultural factors play a role in the way that families react to and cope with major mental illness in a relative, and this moves the understanding of EE into an interactional framework.

During direct observation of high EE family members interacting to solve a problem, the emergence of subclinical psychopathology in the patient triggered responses by the patient's relatives that might be viewed as incipient signs of high EE. Negative interactional cycles escalated, with efforts by the family to squelch the patient's deviant behavior, leading to more intense symptomatic behavior and greater negativity in the cycles. There is evidence also that enduring personality, attributional, and information processing factors in high EE relatives may contribute to the association between EE and relapses in psychotic disorders. For example, families who perceive the causes of the patients' problems or symptoms as being controllable by the patients (i.e., families with high internal locus of control) are more likely to be high EE. On the other hand, families that are rated as having higher tolerance, flexibility, empathy, and communication are more likely to be low EE.

The role of EE as a stressor in schizophrenia and other psychotic conditions is elucidated by the stress–vulnerability–protective factors model. Certain characteristics of relatives (e.g., attributional style) may predispose them to experience stress when they are faced with abnormal, unpredictable, and even dangerous behavior in their ill family member. They attempt to cope with this stress by using any means at hand – even criticism, hostility, or emotional overinvolvement – and this has counterproductive effects on the ill relative, causing that person stress and leading eventually to relapse after many cycles of stress superimposed on vulnerability in both the patient and relatives. The interactional nature of EE that develops in this synchronous fashion does indeed link stress, vulnerability, and protective factors in both the family members and patient.

Treatment as Protection against Stress and Psychosis

The Role of Antipsychotic Medication in Reducing Psychosis

In most, if not all studies, antipsychotic medications have repeatedly been shown to be effective for the treatment of psychosis and the prevention of relapse in schizophrenia. Brown and Birley reported that a

larger percentage of patients who relapsed while taking an antipsychotic medication had experienced a life event compared to patients who relapsed while not on medication. Additional work by Brown and Birley in 1968 also suggested that "life events and reducing or stopping phenothiazines [antipsychotic medication] contribute as precipitants of acute schizophrenia." Similar results were published showing that a significantly greater proportion of the patients who relapsed on medication had experienced an independent life event, compared to nonrelapsing patients not on active medication. Leff proposed that regular antipsychotic maintenance medication produces a prophylactic effect against stress by raising a patient's threshold to relapse; this increased threshold makes the patient less likely to relapse unless exposed to major life stressors.

In a 1992 prospective study, Ventura et al. found that patients who had a psychotic exacerbation while on medication had more independent events in the prior month than did patients who had an exacerbation while off medication. As previously proposed by Leff, Ventura et al. concluded that taking an antipsychotic medication is protective and raises the threshold for relapse. However, another well-designed study did not completely support the hypothesis that life events are more relevant to relapse in patients on maintenance medication.

Behavioral Family Management for Reducing Stress Inherent in High Expressed Emotion

As shown in Figure 2, the daily hassles and challenging or stressful events of everyday life in a family with a relative who has a disabling psychosis can be dealt with in either of two ways. If the family is equipped with good coping and communication skills, these daily problems can be decompressed and solved without the emotional temperature in the family rising above the ill relative's vulnerability threshold, decreasing the risk of relapse. However, families that lack the know-how and skills to cope with the stressors of everyday life can, all too readily, succumb to a cycle of stress, poor communication, and mounting problems, leading to relapse. Given the large proportion of families that show high EE in the United States and most developed countries, an intervention program is essential to cool the family emotional temperature.

Behavioral family management, initially designed by Liberman, Falloon, and their colleagues at the UCLA Center for Research on Treatment and Rehabilitation of Psychosis, comprises education of the family and patient about the nature of the mental illness affecting the patient, information and advice about using community-based resources for treatment and rehabilitation, planned and structured

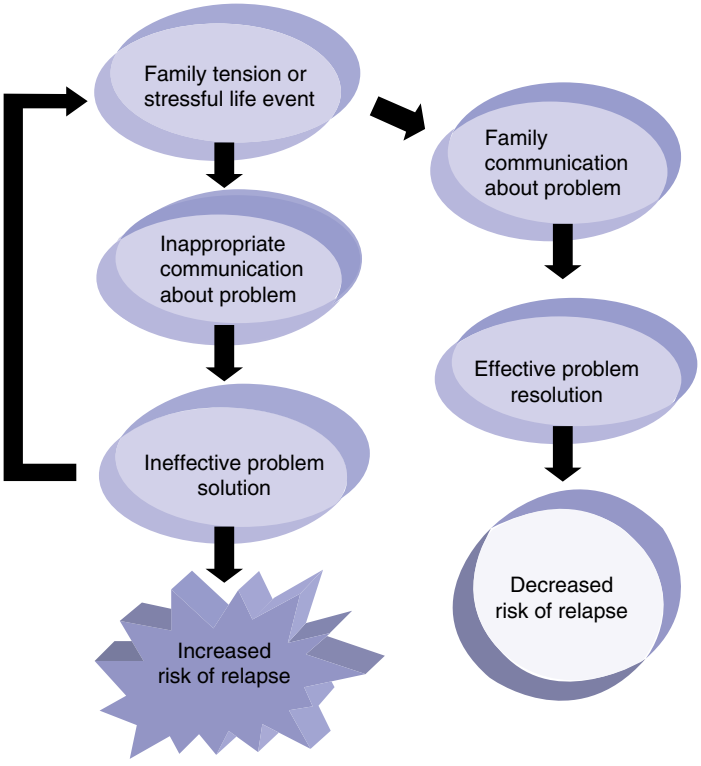

Family tension or stressful life event
Inappropriate communication about problem
Ineffective problem solution
Increased risk of relapse
Family communication about problem
Effective problem resolution
Decreased risk of relapse

Figure 2 Role of the family in preventing relapse.

training in communication skills, and training in problem solving. Research conducted and replicated on behavioral family management and related treatments has revealed that family-based stress management and skills training is a highly effective addition to the long-term clinical management of psychotic disorders. This modality of treatment has shown superiority in reducing relapse and rehospitalization rates, improving rates of remission of symptoms, enhancing social outcomes, and increasing family and economic benefits. Moreover, the technique can be adapted to be useful in a wide variety of settings, employed by many different disciplines of mental health workers, and effective in a spectrum of patient populations.

Coping with Stressful Life Events

Stressful life events occur often enough in the lives of schizophrenia patients to significantly increase the risk of psychotic relapse. Some research suggested that individuals who experience stress, but who do not relapse, possess psychological characteristics and cognitive functioning that may be associated with effective coping. In particular, the use of problem-focused coping strategies has been associated with good neurocognitive functioning and has been shown to lessen the impact of stressful life events by lowering the risk of relapse. High levels of self-esteem, hope, and insight and the perception of available social support were found in schizophrenia patients who reported using active problem-focused coping strategies. Protective mechanisms, such as effective coping behaviors, in patients and their families may have changed stressful situations into more minor events and reduced the risk of relapse. Positive symptoms have been reduced in patients who consistently applied coping strategies to stressful life events and used problem-solving techniques. These findings support the importance of coping behavior as a protective factor and mediator in the vulnerability–stress–protective factors model.

Despite good evidence for the value of using specific problem-focused coping strategies, schizophrenia patients may fail to use them. This differential use of coping strategies may explain the variability of schizophrenia patients' responses to stressful events. Even after training, they may continue to use ineffective strategies and report that the strategies they used most often were the least effective. Despite the fact that successful coping behaviors may be protective and can be learned, relatively few patients seem to avail themselves of the benefits of using coping strategies. In addition, because many schizophrenia patients have aversive reactions to stressful events, they may respond with avoidance and withdrawal to reduce their acute stress. Thus, further longitudinal research is needed to determine whether short-term or brief avoidance could in some way be adaptive or always leads to problems later.

Further Reading

Brown, G. W. and Birley, J. L. T. (1968). Crisis and life change and the onset of schizophrenia. *Journal of Health and Social Behavior* **9**, 203–214.

Bustillo, J., Lauriello, J., Horan, W., et al. (2001). The psychosocial treatment of schizophrenia: an update. *American Journal of Psychiatry* **158**(2), 163–175.

DeLisi, L., Cohen, T. and Maurizio, A. (2004). Hospitalized psychiatric patients view the World Trade Center disaster. *Psychiatry Research* **129**(2), 201–207.

Myin-Germeys, I., Krabbendam, L., Jolles, J., et al. (2002). Are cognitive impairments associated with sensitivity to stress in schizophrenia?: an experience sampling study. *American Journal of Psychiatry* **159**(3), 443–449.

Mylin-Germeys, I., van Os, J., Schwartz, J., et al. (2001). Emotional reactivity to daily life stress in psychosis. *Archives of General Psychiatry* **58**(12), 1137–1144.

Ventura, J., Nuechterlein, K. H., Subotnik, K. L., et al. (2000). Life events can trigger depressive symptoms in early schizophrenia. *Journal of Abnormal Psychology* **109**, 139–144.

Q

Quality of Life

S M Skevington
University of Bath, Bath, UK

© 2007 Elsevier Inc. All rights reserved.

The Concept of Quality of Life
Assessment of Quality of Life
Dimensions of Quality of Life

Glossary

Subjective	Personal.
Objective	Detached from personal views.
Observable	Visible.
Perceptions	Views.
Functioning	What a person is capable of doing.
Encephalograph (EEG)	Measured electrical changes in brain waves.
Eustress	Positive or "good" stress.

The Concept of Quality of Life

Quality of life (QoL) has become a key goal of contemporary health care. It is often confused with standard-of-living. However, standard-of-living refers to the possession of wealth or material goods. Although a certain number of square feet of living space in United States, or oxen owned in Ethiopia, makes a tangible difference in people's lives, it does not necessarily bring greater happiness or well-being; we can observe this in lottery winners. International figures on annual income per capita collected by the World Bank show that once income exceeds a critical level – $13,000 in 1995 – the close relationship between subjective well-being and income becomes increasingly loose and dissipated. Furthermore, in the world's richest countries standard-of-living increases do not appear to make a significant difference to people's QoL, a finding that is highly perplexing for policy makers.

A consensus about how to define and measure QoL is still widely debated. In the 1970s, its definitions contained terminological similarities to definitions of stress. At a time when stress was seen as a phenomenon that exceeded people's resources, in order to provide a good QoL these resources needed to be adequate in terms of satisfying people's wants, needs, and capacities. Since then, definitions have placed greater emphasis on people's subjective perceptions of the important features of their life and, in particular, explored the varied meanings ascribed to these experiences. The ways in which people interpret life's events (e.g., as stressful or pleasant) affects how they see their QoL. Wenger et al. (1984) defined QOL as "an individual's perceptions of his or her functioning and well-being in different domains of life." Judgments about QoL are now seen as a rich interplay and balance between how people see their internal state, such as the tension in muscles or happiness, and the external events that impinge on them from their environment, such as changing jobs or being bereaved.

Assessment of Quality of Life

The history of QoL assessment measures shows that some of the earliest versions were designed by clinicians for their own use in distributing health-care resources, for example the assessment of Quality Adjusted Life Years (QALYS). This approach excluded a systematic assessment of patients' views, with the assumption that clinicians held a reasonably accurate view of their patients' QoL. However, research has consistently shown that QoL judgments by clinicians and other observers (e.g., spouses or caregivers) often correlate weakly with the views held by the patients about their own QoL. Such proxy information is therefore unreliable, and this is particularly apparent in the case of dementia or stroke, in which the patients' own perceptions can be difficult or impossible to obtain. This discrepancy arises from the observation that many aspects of QoL are simply not open to inspection but are hidden from view like the submerged parts of an iceberg. Furthermore, when measures have limited their contents to subjective reports of observable behaviors such as mobility and

activities of daily living, it has resulted in the assessment of a constricted rather than holistic concept of QoL, due to the exclusion of important but invisible QoL issues such as spirituality and social relationships.

The need for QoL assessment in health has been driven by the problem-solving demands of clinical practice. The aim was to better understand the patient's priorities in order to deliver appropriate care. However, this approach may have provided an unduly negative view of QoL. Patients expect to report their problems during a consultation and to do this succinctly. In these time-limited conditions, they are unlikely to also discuss the positive qualities of life such as the pleasure they get from playing with children, listening to Mozart, or seeing a dewdrop on a leaf. Consequently, many measures have tended to focus on assessing symptoms and dysfunctional aspects of life to the exclusion of the good life. This not only gives health professionals the view that the QoL of their patients is worse than it is, but it also causes the patients who complete the questionnaire to report feeling miserable after focusing unduly on the bad things instead of a more balanced life. Even highly disabled, chronic-pain patients entering a pain management program after many surgical and pharmacological treatments report that chronic pain affects only half the important qualities of life when a range of positive and negative issues are presented for evaluation. The good life was acknowledged by Aristotle and several eighteenth-century philosophers, but has only recently been endorsed again as pertinent to outcomes measurements in health care. Hence, more recent definitions of QoL have sought to emphasize a positive orientation: "The essential characteristics of life, which in the general public is often interpreted as the positive values of life, or the good parts of life, or the total existence of an individual, group or society" (Lindstrom, 1992).

The Example of Sleep as a Quality of Life

Another debate concerns the type of information that constitutes the best measure of QoL, and here sleep provides a suitable example. Researchers have assessed sleep-related QoL by tracing electroencephalograph (EEG) waves, considering this to be an objective indicator. Others assess the timing of sleep and wakening, insomnia patterns, medication use, and so forth, in looking at people's objective perceptions of their sleep. However, we can argue that information derived from either of these channels of inquiry tells us only indirectly about people's QoL. These types of information require further interpretation by the researchers, and their own assumptions may lead them to draw an erroneous conclusion. Would the people you know see your QoL as good or bad if you told them you slept 3 h a night? What about 12 h a night? Without additional information about whether that sleep was refreshing, it is not possible to draw an accurate conclusion. For this reason the most reliable way to avoid this pitfall is to ask subjective QoL questions such as whether people find their sleep to be refreshing or not. Conclusions about these qualities cannot, therefore, be deduced; direct questioning about the meaning of events is required. Like stressful life events (e.g., divorce), it is the interpretation that matters. Although all these three types of information are themselves valuable and can be used together to form a fuller picture, we contend that QoL is best measured as a subjective phenomenon and that the use of the term should be restricted to this type of information.

The World Health Organisation Definition

Questions have been raised about whether QoL is a global phenomenon or just a Western concept. Although culture has sometimes been included in QoL models, it is usually seen as extraneous and additional, rather than taking a central role in explaining the diverse ways in which different cultural groups interpret QoL. Concerned to understand QoL relating to global health, the World Health Organisation views culture as a "lens" through which people interpret the experiences that affect their QoL. It defines QoL as "an individual's perception of their position in life, in the context of the culture and value systems in which they live, and in relation to their goals, expectations, standards, and concerns" (World Health Organisation Quality of Life Group, 1994: 43). This unusual definition also implies that comparisons are used in helping to make decisions about QoL; that in evaluating their own QoL, people consider how far they have achieved their goals, expectations, and standards, in addition to considering problematic concerns.

Dimensions of Quality of Life

So what does a good QoL consist of, and is there consensus about whether these dimensions are universal? Many hundreds of QoL questionnaires have included the two broad components of physical and mental health identified by Descartes as essential. However, physical health is operationalized in some assessments in terms of symptoms; adding symptoms does not give a clear view of people's QoL because, paradoxically, patients with many symptoms often report a relatively good QoL and, conversely, those

with very few symptoms report it to be poor. Measures also often include estimates of independence or functional status, such as mobility, activities of daily living, and so on. Physicians have traditionally favored those QoL dimensions for which they can independently inspect corroborating evidence alongside the subjective reports. Some issues such as spirituality have largely been ignored as being too difficult to measure and/or irrelevant to health and health care.

The universality of QoL and its dimensions have been addressed by the World Health Organisation Quality of Life (WHOQOL) Group, which has confirmed that physical, psychological, independence, social relations, environmental and spiritual dimensions are important in a wide range of cultures and settings and to sick and well people. Within these broad domains, the group confirms 25 detailed aspects (facets) of QoL that are reliable and valid internationally. A series of novel procedures was used by the WHOQOL Group to create several multilingual instruments and a short version – the WHOQOL-Bref – is now available in over 50 different languages. In addition to the main measures, add-on modules of extra questions have recently become available to support work on spirituality, human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS), older adults, and pain. The WHOQOL results provide support for the view that there is broad recognition and universal agreement about a concept called QoL, even though it may not be labeled as such, and that there are cultural variations in the relative importance of the dimensions included, as well as in how these issues themselves are expressed.

Uses of Quality of Life Assessments

QoL assessments are used for monitoring outcomes (e.g., after policy change). In particular, they are valued widely as a suitable outcome measure in randomized controlled clinical trials. Although outcomes are the main focus here, it is worth noting that QoL can also be investigated as a process. Because psychometric measures are designed to tap into outcomes, the results from these instruments provide merely a snapshot of QoL at one time in an ever-adjusting process. Different theories and techniques may be needed to investigate how and when QoL changes over time and precisely which factors (e.g., stress and eustress) affect positive and negative changes.

Further Reading

Coen, R. F. (1999). Individual quality of life and assessment by carers or 'proxy' respondents. In: Joyce, C. R. B., O'Boyle, C. A. & McGee, H. (eds.) *Individual quality of life: approaches to conceptualization and assessment*, pp. 185–196. Amsterdam: Harwood Academic.

Inglehardt, R. and Klingemann, H. D. (2000). Genes, culture, democracy and happiness. In: Diener, E. & Suh, E. M. (eds.) *Culture and subjective well-being*, pp. 185–218. Cambridge, MA: MIT Press.

Layard, R. (2005). *Happiness: lessons from a new science*. London: Penguin Books.

Lindstrom, B. (1992). Quality of life – a model for evaluating health for all: conceptual considerations and policy implications. *Sozial und Preventivmedizin* **37**, 301–306.

Shin, D. C. and Johnson, D. M. (1978). Avowed happiness as an overall assessment of quality of life. *Social Indicators Research* **5**, 475–492.

Skevington, S. M. (2002). Advancing cross-cultural research on quality of life: observations drawn from the WHOQOL development. *Quality of Life Research* **11**, 135–144.

Skevington, S. M., Lotfy, M. and O'Connell, K. A. (2004). The World Health Organisation's WHOQOL-BREF Quality of Life assessment: psychometric properties and results of the international field trial – a report from the WHOQOL Group. *Quality of Life Research* **13**, 299–310.

Skevington, S. M., Sartorius, N., Amir, M., et al. (2004). Developing methods for assessing quality of life in different cultural settings: the history of the WHOQOL instruments. *Social Psychiatry and Psychiatric Epidemiology* **39**, 1–8.

Skevington, S. M. and Wright, A. (2001). Changes in the quality of life of patients receiving anti-depressant medication in primary care: validating the WHOQOL-100. *British Journal of Psychiatry* **178**, 261–267.

Staquet, M. J., Hayes, R. D. and Fayers, P. M. (1998). *Quality of life assessment in clinical trials: methods and practice*. Oxford: Oxford University Press.

Wenger, N. K., Mattson, M. E., Furberg, C. D., et al. (1984). Assessment of quality of life in clinical trials of cardiovascular therapies. *American Journal of Cardiology* **54**, 908–913.

World Health Organisation Quality of Life Group (1994). The development of the World Health Organisation Quality of Life assessment instrument (The WHOQOL). In: Orley, J. & Kuyken, W. (eds.). *Quality of life assessment: international perspectives*, pp. 41–60. Berlin: Springer-Verlag.

World Health Organisation Quality of Life Group (1995). The World Health Organisation Quality of Life assessment (WHOQOL): position paper from the World Health Organisation. *Social Science & Medicine* **41**, 1403–1409.

World Health Organisation Quality of Life Group (1998). The World Health Organization Quality of Life assessment (WHOQOL): development and general psychometric properties. *Social Science & Medicine* **46**, 1569–1585.

R

Racial Harassment/Discrimination

D R Williams and S A Mohammed
University of Michigan Ann Arbor, MI, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction: Discrimination Persists
Studies of Discrimination and Health
Understanding the Context of Discrimination and Responses to It
Measurement Issues
Pathways from Discrimination to Health
Conclusion

Introduction: Discrimination Persists

Discrimination and Racial Harassment

Several lines of evidence converge to suggest that discrimination is an important feature of life in race-conscious societies in which race is an important determinant of access to a broad range of societal benefits and rewards. First, social psychological theory identifies discrimination as a key aspect of intergroup relationships that serve to reinforce the symbolic boundaries that separate one social group from another. Second, there is broad recognition that discrimination persists for socially disadvantaged populations. The American public acknowledges that a high level of discrimination and prejudice exists. National public opinion data reveal that most Americans concede that racial minorities suffer from a fairly high level of racial bias. Blacks are perceived as experiencing the most discrimination, followed by Hispanics and American Indians, and then Asians. They also perceive that some White ethnic/religious groups such as Jews, Italians, and Catholics experience more discrimination than Whites in general. In one national study, for example, two-thirds of the American population reported that blacks suffer a lot or a tremendous amount of discrimination.

Some of the best evidence of the persistence of discrimination comes from audit studies in which carefully matched Black and White applicants with identical qualifications apply for jobs or housing. Audit studies in employment, for example, find that discrimination favors the White over a Black applicant in one in five audits. Other audit studies suggest that discrimination persists for Blacks and other minorities in the purchase of a broad range of goods and services. For example, one carefully executed study found racial differences in the best price offered for the purchase of a new car. In this study, Black and White testers were sent to automobile dealerships to negotiate the price of a new car. Although all the testers were carefully trained to follow the same script, there were large racial and gender differences in the final price offered for the new car, with White males offered the lowest prices, followed by White females, African American males, and African American females. Similarly, a recent report from the Institute of Medicine documented systematic and pervasive discrimination in the quality and intensity of medical treatment received by Blacks and other minorities in the United States. Thus, there is abundant evidence that discrimination persists in contemporary society. However, not all experiences of discrimination are visible to their intended target. Frequently, victims of discrimination lack all of the details about an interpersonal transaction that they need to know in order to establish that discrimination has occurred. Although self-reports of bias understate the full extent of exposure to discrimination, they capture an important part of discriminatory experiences that are likely to be consequential for health.

Discrimination as a Stressor

This article focuses on aspects of discrimination that are perceived by individuals in society. Subjective experiences of discrimination or unfair treatment are viewed as an important type of stressful life experience that many traditional measures of stress fail to assess. Equity theory suggests that perceptions of unfair treatment can lead to negative emotional reactions and psychosomatic symptoms. Discrimination

can also be stressful because it can reinforce negative social stigma and highlight low social regard. Qualitative descriptions of experiences of discrimination indicate that they can induce considerable levels of psychological distress. Laboratory studies have examined the experimental manipulation of unfair treatment and have found that it leads to elevated levels of psychological distress for a broad range of social groups. Other studies have directly measured physiological responses to racially stressful material in the laboratory setting. These studies have used mental imagery and videotaped vignettes of discriminatory behavior. Most, but not all, of these studies have found that exposure to discrimination leads to cardiovascular and psychological reactivity. An important methodological strength of laboratory studies is that they do not rely on respondents' self-reports of discrimination. However, observed effects in the laboratory are short term, making the extent to which acute physiological arousal under laboratory conditions generalizes to chronic elevation of stress processes and mechanisms in real life unclear. Some recent research documents that high reactivity to acute stressors in laboratory settings is predictive of elevated rates of hypertension several years later. Most of the early laboratory studies focused on Blacks. It is not clear from these studies, however, whether the physiological reactivity to racism-linked experiences is unique or enhanced in comparison to reactivity to other types of stressors.

Studies of Discrimination and Health

Mental health status is the most frequently used health measure in studies of discrimination. Multiple mental health indicators have been used. These include measures of well-being, self-esteem, perceptions of control, psychological distress, anger, and specific psychiatric disorders such as major depression, generalized anxiety, and substance use. These studies have found an inverse relationship between discrimination and mental health – higher levels of discrimination were associated with poorer mental health status. Most of the early research studies in this area were based in the United States and tended to focus on African Americans. Recent studies have documented similar patterns of associations for other minority groups in the United States. For example, studies of Mexican Americans, Chinese Americans, and American Indians have found that perceived discrimination is associated with higher levels of psychological distress. A growing number of international studies documented similar patterns. Studies of Asian immigrants in Canada; Iranians, Turks, and Moroccans in the Netherlands; immigrants in Finland; Blacks, Asians, and Arabs in Ireland; and multiple ethnic minority populations in the United Kingdom found that discrimination is positively associated with psychological distress. Ongoing research is also exploring these relationships in Australia, Brazil, New Zealand, and South Africa.

Measures of perceived discrimination have also been studied in relationship to multiple measures of physical health. Studies using a single-item global self-rated health measure and other self-ratings of health or checklists of chronic illnesses have also fairly consistently documented that discrimination is related to poorer health status. Two studies found that discrimination makes an incremental contribution above socioeconomic status (SES) in explaining Black-White differences in self-reported health status. Several studies have also examined the association between discrimination and blood pressure levels or hypertension. Some studies found a clear positive association between discrimination and elevated blood pressure, whereas others found that this effect varies with the coping style, sex, SES, or race. Still other studies found no association between discrimination and blood pressure.

Recent research has also paid attention to a broader range of health outcomes. Whereas perceptions of discrimination were unrelated to self-reported heart disease in some early studies, recent research documented a positive association between chronic discrimination and the onset of subclinical heart disease (intima media thickness and coronary calcification). Recent studies also found that maternal exposure to discrimination was associated with preterm deliver and low birth weight. Several studies have also examined the association between perceived racial bias and health behaviors. They have found a positive association between discrimination and cigarette use and alcohol consumption.

Understanding the Context of Discrimination and Responses to It

Some of the inconsistency in findings in studies of discrimination and health could be due to inadequate attention to characterizing the social and psychological context in which discrimination occurs as well as the specific characteristics of the discriminatory incident. First, personality characteristics could predict variations in the association between discrimination and health. Beliefs about the self, such as self-esteem, mastery, racial consciousness, and identity might affect an individual's appraisal and thus the potential threat of a discriminatory experience. Second, contextual variables such as social support could also enhance an individual's capacity to cope and respond

to discriminatory experiences. At the present time, we do not have a clear understanding of the contextual factors that can foster either vulnerability or resilience to perceived bias.

Third, specific characteristics of the discriminatory experience and its particular context can be important determinants of its impact. Key aspects of discriminatory experiences include the type or domain in which it occurs, the magnitude of the incident, its temporal characteristics (acute versus chronic), and the nature of the relationship between this stressor and other race-related and non-race-related stressors. Inadequate attention has also been given to the extent to which the effects of discrimination may vary by the social characteristics of the perpetrator. For example, some limited evidence suggests that the negative health impact of perceived discrimination on African Americans is greater when the perpetrator is also African American than when the perpetrator is White.

Fourth, it is also likely that there are complex interactions between exposure to discrimination and particular coping strategies or styles. In addition to examining traditional problem-focused and emotion-focused measures of coping from the stress literature, research on discrimination needs to attend to discrimination-related measures. Heightened vigilance is one such measure. Because racism is often deeply embedded in the structure and culture of society, it can pose an ever-present threat for socially stigmatized groups. Some individuals respond to this by engaging in psychological and behavioral anticipatory coping strategies to mitigate the perceived dangers in the environment. This can lead to heightened physiological arousal that can adversely affect health. Using denial as a coping strategy is a theoretically important but difficult concept to measure. The literature suggests that, because experiences of discrimination can pose threats to the self, some individuals respond to its occurrence by minimizing or even denying its occurrence. There is also the suggestion in the literature that denying the occurrence of bias is damaging to health.

Measurement Issues

A prerequisite for understanding the association between discrimination and health is the development of valid and reliable measures that fully characterize exposure to discrimination. The literature identifies several issues that affect the appropriate assessment of discrimination.

Perceived Discrimination Is Subjective

An important issue in the assessment of discrimination is the challenge of obtaining valid and reliable indicators of subjective phenomena such as experiences of discrimination. It is possible that the salience of racial identity or perceived racism could lead some minority group members to perceive incidents as racist that may not be or, worse, could lead to the development of a mind-set in which they perceive experiences to be products of racism when in fact they were not. In fact, some limited evidence suggests that the use of discrimination terminology, in which race or gender is made salient in the assessment of discrimination, can induce respondent biases. This work emphasizes that questions with embedded terminology require both a description of an experience as well as an interpretation of that experience. Moreover, in the cooperative context of a research interview a respondent may make inferences about the questioner's intent and interests that could lead the respondent to try to provide information that is consistent with the questioner's perceived intent. All of this can lead to interpretive response bias that could produce higher estimates of discrimination, driven by the specific terminology used. Thus, questionnaires that ask repeated questions about "racial discrimination" or experiences "because of your race" could produce demand characteristics in which the respondent believes that it is desirable to the interviewer to report such experiences. This could lead to overreports of discrimination.

One promising approach in measuring discrimination seeks to reduce the salience of race by first asking respondents whether they have been treated "unfairly" in multiple domains of life. After respondents have endorsed an experience of unfair treatment, they are asked to attribute a reason. Potential reasons include race and ethnicity, but also allow for other factors such as gender, age, religion, weight, and sexual orientation. This approach enables respondents to report on all instances of unfair treatment, but allows the researcher to separate instances attributed to race from those linked to other reasons. Relatedly, this measurement approach allows for the evaluation of the extent that discrimination based on race is more deleterious to health than discrimination based on other social factors. Inadequate attention has been given to the relative impact of racial discrimination compared to other types of discrimination in the extant research to draw firm conclusions. However, some studies suggest that it is the generic perception of unfair treatment that is deleterious to health, irrespective of the specific attribution regarding the cause of the incident. Research is needed to identify the optimal approaches to the measurement of racial discrimination.

The available evidence suggests that concerns about the overreporting of discrimination may be

overblown. First, making race salient in the assessment of discrimination can lead to underreporting of discrimination. Respondents vary in their thresholds of what constitutes discrimination and may fail to report as discriminatory incidents that were not perceived as very serious. Second, respondents appear to interpret the concept of discrimination as intended by researchers and self-reports of discrimination are consistent with objective experiences. Third, two studies have documented that baseline mental health status (psychological distress and major depression) is unrelated to subsequent reports of discrimination. Fourth, because reporting discrimination can negatively affect self-esteem and perceptions of control, at least some stigmatized group members are likely to minimize and deny experiences of discrimination. Thus, underreporting is as important a threat to the validity of self-reports of discrimination as overreporting. Nonetheless, it is important for researchers to adjust reports of discrimination for potentially confounding psychological characteristics such as social desirability, neuroticism, and other indicators of negative affect. Such adjustment strengthens the methodological rigor of a study and increases the likelihood that observed associations between perceived discrimination and health are not distorted by underlying psychological factors. Including such controls is especially important when the measure of health status is also based on self-report.

Several strategies that have been used to improve the accuracy of the reporting of stressors should also be applied to the study of discrimination. These include the use of cues to memory (such as visual representations and reminders of personally salient events), wording the questions in ways that clearly indicate the domain of the experience being captured, and using a life-events calendar to facilitate dating the onset and resolution of experiences of discrimination.

Perceived Discrimination Is Multidimensional

Capturing the exposure to discrimination in its full multidimensional complexity and assessing the cumulative burden of such exposure over the life course is also another major challenge. Some early studies of discrimination used single-item measures. Such assessment understates the actual level of discrimination. However, many current studies assess exposure to discrimination using retrospective recall to capture exposure within a 30-day, 1-year, 3-year or lifetime time frame.

Comprehensively capturing discrimination requires its assessment in multiple areas of social life. Like other stressors, discrimination is multidimensional and should be assessed in all relevant domains. The stress literature indicates that the most commonly measured types of stressors include life events, chronic stress, and daily hassles. Life events are major discrete stressors that are readily observable. Daily hassles are minor but often chronic irritations. Chronic stressors are ongoing problems that are typically role-related. There are analogues to all of these types of stressors in current measures of discrimination. Major acute experiences of racial bias are the most commonly assessed type of discriminatory experience. The literature on stress indicates that chronic stressors may be stronger predictors of the onset and course of illness than acute life events. Scales such as the Everyday Discrimination Scale attempt to capture persistent and recurring everyday chronic minor experiences. Future research needs to give more attention to capturing the frequency and duration of chronic discrimination in multiple social contexts, such as employment, educational, and public settings.

Other major types of stressors include traumas, nonevents, and macro-stressors. Developing analogs of these in the assessment of discrimination is a promising direction for future research. Traumas are acute stressors that are severe and overwhelming in impact, such as being kidnapped or raped. Macro-stressors are large-scale systems-related stressors such as economic recessions. Nonevents are desired and expected experiences that fail to occur. The stress literature indicates that the various types of stressors can have independent effects on health. Therefore, assessing all relevant types of stressors is a perquisite to an assessment of the full impact of stress on health.

A small but growing body of research on historical trauma and its intergenerational effects on the health of American Indians illustrates the benefit of expanding our definitions of race-related stressors. Historical trauma is viewed as a cumulative and collective psychological wounding over the life span and across generations as the result of the history of genocide and assimilation that American Indians experienced at the hands of Europeans. Studying historical trauma among American Indians is not unlike existing studies of other generational group traumas, such as studies of the effects of the Jewish Holocaust on the physical and mental health, as well as the social, economic, and political contexts, of descendants of Holocaust survivors. It has been argued that historical trauma contributes to a host of issues in the American Indian population, including unresolved grief, alcoholism, depression, anxiety, high rates of suicide, homicide, problematic gambling behaviors, domestic violence, child abuse, poverty, low education levels, and various physical diseases. The literature recognizes that although the concept of historical trauma is theoretically intuitive, there are challenges

in distinguishing historical versus contemporary experiences, understanding how historical trauma is imparted across generations, and defining and measuring the levels of historical trauma. However, scales with good psychometric properties have been developed to assess historical trauma and find that as many as one-half of American Indians think regularly about these historical losses. Moreover, recent empirical studies have found support for an inverse association between historical trauma and health. Clinical interventions to address historical trauma have also been developed. A promising area of research would be to examine the generalizability of this pattern across health outcomes and American Indian cultures, as well as, seeing whether other highly visible traumatic historical events have similar effects for other groups.

Capturing Life Course Exposure

There are important measurement challenges in terms of identifying the appropriate questions that could capture the duration and frequency of discrimination over the life course. An isolated experience of discrimination, just like a random stressor, is unlikely to have long-term negative effects on health. The literature suggests that the key issue is being able to identify discriminatory experiences that are enduring and chronic. A life course perspective is critical in assessing the cumulative burden of discrimination over an individual's lifetime. Measures that do an excellent job of capturing exposure to all of the relevant types of discrimination do not currently exist. Overcoming the challenges linked to declines in the accuracy of recall with the passage of time will not be easy, but this is crucial to obtaining better measures of lifetime exposure to discrimination. Such data are necessary to begin to identify the lag times between exposure to discriminatory stressors and adverse changes in health. Life course assessment of discrimination must also be coupled with an equally comprehensive measurement of other types of stressors. The present-day literature does not provide a clear understanding of how experiences of discrimination relate to other stressors and combine, in additive and interactive ways, to affect specific health outcomes and trajectories of health status over time.

Attributional Ambiguity

Perceived discrimination is based on the perception of the individual. The subjective nature of discrimination and the ambiguity that occurs in much interpersonal interaction can lead to uncertainty regarding the attribution of specific incidents of bias or unfair treatment. It is possible that having to make sense of ambiguous social interactions can be physiologically stressful. The worry and rumination regarding the causes of experiences of discrimination may be an added burden that can be consequential for the health of nondominant group members. It is thus possible that the degree of ambiguity attending the perception of a discriminatory experience could negatively affect health. This issue deserves careful examination in future research.

Pathways from Discrimination to Health

Researchers have generally found a weaker association between discrimination and health for measures of physical health than for measures of mental health. Physical health outcomes typically assessed are chronic conditions involving complex etiological factors and varying patterns of development and progression over long periods of time. Greater attention needs to be given to the hypothesized pathways by which discrimination might affect health and to identifying the appropriate points along that pathway at which the individual is more vulnerable to exposure to discrimination and other stressors.

The stress literature reveals that one way that stressors can affect health is by giving rise to negative emotional states, such as anxiety and depression, which in turn can directly impact biological processes or patterns of behavior that affect disease risk. Accordingly, one of the pathways by which discrimination can affect physical health outcomes may be indirectly, through psychological distress. Thus, discrimination may lead to elevated psychological symptoms, which, in turn, may lead to chronic physiological arousal. In addition, the negative emotional states created by experiences of discrimination might lead to health behaviors that may ultimately affect disease risk, such as impaired sleep patterns, decreased physical activity, increased substance use, and consumption of more food than usual. It was noted earlier that prior research indicates that discrimination is positively associated with alcohol consumption and cigarette smoking. These and other health-related behaviors may be a pathway by which perceived bias has health consequences. The stress of discrimination and the negative emotional states created by it could also lead to lower levels of adherence to medical recommendations. This latter mechanism has not yet been explored in the literature. The potential effects of discrimination and other stressors on medical compliance emphasize the need to assess the contribution of discrimination not only to the onset of disease but also to its severity and course.

Our understanding is currently limited regarding how exposure to discrimination leads to changes in particular biological responses and health behaviors. Further research is needed to identify the

physiological mediators of the effects of discrimination so that the specific biological pathways by which discrimination can adversely affect health may be studied. Such research should assess the conditions under which specific physiological systems, such as the cardiovascular, neuroendocrine, and immune systems, are affected by particular types of discrimination. We are presently unaware of the genetic and psychological factors that can lead some organ systems to be more vulnerable than others to the effects of discrimination on health.

Conclusion

Scientific research indicates that perceived discrimination is associated with poor health status, with the association being strongest for mental health. However, existing evidence does not clearly indicate the extent to which exposure to perceived discrimination leads to increased risk of disease, the conditions under which this is more or less likely to occur, or the underlying mechanisms and processes linking this stressor to health status. Advances have been made in the conceptualization and measurement of discrimination, but there is still more work to be done. The most urgent need is for the theoretical identification and the empirical verification of the plausible pathways by which experiences of bias can affect various health outcomes. At the same time, this field of research is clearly moving from infancy to childhood. The literature has now identified major gaps in this area, as well as promising conceptual, methodological, and analytic tools that are needed for the rigorous examination of the association between perceived discrimination and multiple indicators of health.

See Also the Following Articles

Cultural Factors in Stress; Refugees, Stress in; Ethnicity, Mental Health.

Further Reading

Blank, R. M., Dabady, M. and Citro, C. F. (eds.) (2004). *Measuring racial discrimination: panel on methods for assessing discrimination.* National Research Council Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Clark, R., Anderson, N. B., Clark, V. R., et al. (1999). Racism as a stressor for African Americans: a biopsychosocial model. *American Psychologist* **54**(10), 805–816.

Dion, K. J. (2001). The social psychology of perceived prejudice and discrimination. *Canadian Psychology* **43**(1), 1–10.

Fix, M. and Struyk, R. J. (1993). *Clear and convincing evidence: measurement of discrimination in America.* Washington, DC: Urban Institute Press.

Gomez, J. P. and Trierweiler, S. (2001). Does discrimination terminology create response bias in questionnaire studies of discrimination? *Personality and Social Psychology Bulletin* **27**(5), 630–638.

Harrell, J., Hall, S. and Taliaferro, J. (2003). Physiological responses to racism and discrimination: an assessment of the evidence. *American Journal of Public Health* **93**(2), 243–248.

Harrell, S. (2000). A multidimensional conceptualization of racism-related stress: implications for the well-being of people of color. *American Journal of Orthopsychiatry* **70**(1), 42–57.

Krieger, N. (1999). Embodying inequality: a review of concepts, measures, and methods for studying health consequences of discrimination. *International Journal of Health Services* **29**(2), 295–352.

Krieger, N., Smith, K., Naishadham, D., et al. (2005). Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health. *Social Science and Medicine* **61**(7), 1576–1596.

Paradies, Y. (2006) A systematic review of empirical research on self-reported racism and health. *International Journal of Epidemiology* **35**, 888–901.

Whitbeck, L. B., Adams, G. W., Hoyt, D. R., et al. (2004). Conceptualizing and measuring historical trauma among American Indian people. *American Journal of Community Psychology* **33**(3–4), 119–130.

Williams, D. R. and Neighbors, H. W. (2002). Racism, discrimination, and hypertension: evidence and needed research. *Ethnicity and Disease* **11**, 800–816.

Williams, D. R., Neighbors, H. W. and Jackson, J. S. (2003). Racial/ethnic discrimination and health: findings from community studies. *American Journal of Public Health* **93**(2), 200–208.

Rape *See:* Sexual Assault.

Receptors, Classical *See:* Membrane Glucocorticoid Receptors; Glucocorticoid Receptor Mutants and Polymorphisms; Corticotropin-Releasing Factor Receptors; Corticosteroid Receptors; Corticotropin Releasing Factor Receptor Deficiency in Mice.

Recovery from Stress

S Sonnentag
University of Konstanz, Konstanz, Germany
C Fritz
Bowling Green State University, Bowling Green, OH, USA

© 2007 Elsevier Inc. All rights reserved.

The Recovery Concept
Relevance of Recovery
Factors Influencing Recovery
Recovery from Job Stress
Conclusion

The Recovery Concept

Research differentiates between two core phenomena related to the stress process: reactivity and recovery. Reactivity refers to the ensuing physiological and psychological changes caused by a stressor and is reflected in the elevation of strain indicators compared to a baseline. Recovery refers to the processes during the poststressor period, when strain indicators return to their baseline levels. Slow recovery from the stressor can be seen as one instantiation of prolonged activity. The duration of the recovery process differs widely among the various indicators. For example, under normal conditions, blood pressure and heart rate recover within a few minutes after the termination of the stressor. However, it can take much longer until neuroendocrine parameters reach the prestressor levels. For example, a literature review based on a total of 77 studies concluded that norepinephrine recovery was complete within 1 h after the termination of the stressor, whereas adrenaline levels often remained elevated, particularly when urinary measures are used. Meta-analytic evidence suggests that salivary cortisol recovery is completed within 40–60 min after the termination of the stressor. Subjective variables such as mood may take hours to return to prestressor levels. Overall, studies show that the time needed for complete recovery largely depends on person-specific and situation-specific factors.

In laboratory research, recovery most often is initiated by a poststressor rest period, during which the stressor is absent. The degree of recovery can be analyzed with several methods, with curve-fitting and related approaches being the most promising and reliable ones, particularly for analyzing cardiovascular recovery. In field research, it is more difficult to identify recovery periods exactly. For example, when examining recovery from job stress we might regard the weekend as the recovery period. However, this perspective assumes that during the weekend no other stressors (e.g., from the family domain) are present.

Relevance of Recovery

Quick recovery after the termination of a stressor is assumed to be an adaptive response pattern, whereas slow recovery is assumed to be maladaptive. Slow recovery implies that stress responses such as elevated blood pressure or increased levels of catecholamines and cortisol persist over longer periods of time and therefore might harm the organism. For example, McEwen argued that a prolonged allostatic response is one of the core processes that contributes to allostatic overload and subsequently might lead to illnesses such as cardiovascular diseases.

Empirical evidence that the recovery process is crucial for health comes for example from a longitudinal study on nurses. Elevated cortisol levels after work (but not during work) predicted health-care costs 5 years later. Pieper and Brosschot summarized empirical studies demonstrating that prolonged cardiovascular recovery after a stressor predicted general heart rate and blood pressure levels 3–6 years later. A recent study with a community sample points in a similar direction. Individuals who showed slow systolic blood pressure recovery from psychological tasks had elevated clinical systolic and diastolic blood pressure levels 3 years later, even after controlling for the initial clinical blood pressure levels and other variables such

as body-mass index, education, and use of tobacco products. However, before we draw any causal conclusions about slow recovery being the cause of elevated clinical blood pressure levels, we must taken into account that cardiovascular risk factors such as race or lack of fitness tend to be related to slow cardiovascular recovery and might cause both slow recovery and high chronic blood pressure.

Factors Influencing Recovery

Research suggests that slow recovery is associated with personal as well as situational factors. For example, individuals with negative emotional dispositions show elevated cardiovascular parameters during poststressor periods. The findings are particularly impressive with respect to hostility. Specifically, hostile persons show elevated levels of systolic blood pressure during sleep. Similarly, in experimental settings that involve the confrontation with a stressor, high hostility is associated with slow systolic blood pressure recovery. In addition to hostility, low self-enhancement (i.e., the tendency of not experiencing highly favorable cognitions about oneself) is related to slow cardiovascular recovery during the poststressor period.

Not only cardiovascular recovery, also cortisol recovery might be hampered by personal factors. Meta-analytical findings show that clinically depressed individuals show higher cortisol levels during the poststressor period than do nondepressed individuals (but not higher baseline or reactivity levels). Moreover, people with low self-esteem show elevated cortisol recovery levels after the experience of failure. In addition, gender and race might also play a role with respect to recovery processes.

With respect to situational factors associated with slow recovery, we can differentiate between acute situational factors present in the specific setting and more chronic situational variables. Acute specific situational factors related to slow recovery refer to the type of stressor and to processes during the poststressor period. Compared with other acute stressors, emotionally involving and anger-provoking situations are associated with a greater delay in cardiovascular and cortisol recovery. It is important to note that slow recovery after the provocation of anger seems to be particularly prominent in men, as opposed to women.

Adrenaline recovery is prolonged after highly demanding situations characterized by a low degree of controllability. With respect to field situations, research suggests that, particularly after periods of high workload and high work intensity, the recovery of adrenaline is delayed.

When it comes to activities that affect recovery during the poststressor period, distraction from the preceding stressor seems to be highly important. Situations that offer distractions from the stressor support faster recovery than situations that allow individuals to ruminate about the stressful event. In addition, music – particularly classical music – helps to reduce blood pressure levels during the poststressor period.

In addition to factors associated with the situation in which the stressor occurs, chronic stressors and other chronic factors may impact recovery processes. People who are exposed to chronic stressors show delayed blood pressure and immunological recovery from acute stressors. Not only chronic stressors but particular stressors in combination with lack of resources are related to slow recovery. For example, in a study with firefighters, it was found that individuals who had experienced major life events during the past 12 months and who received only little social support in various life domains displayed particularly slow cardiovascular recovery. Results from several studies in the work context point in a similar direction. Employees in high-strain jobs (i.e., jobs associated with a high level of demands and low control at the same time) show prolonged cardiovascular and cortisol recovery. However, note that this picture is not consistent across all studies.

Prolonged recovery in people facing stressful life situations is also reflected in research using subjective measures. Individuals in stressful jobs experience a higher need for recovery and feel that it takes them longer to unwind from work. The subjective need for recovery is related to both neuroendocrine parameters and indicators of impaired well-being such as psychosomatic complaints and burnout.

To sum up, personal variables such as hostility or depression are related to slow recovery. Among the most important situational variables that are associated with prolonged recovery are situations that are anger-provoking and emotionally involving and situations characterized by high workload. In addition, chronic background stressors and lack of resources (e.g., low social support) make it more likely that recovery is delayed. With respect to factors that promote recovery, distraction from the stressful situation seems to be highly important.

Recovery from Job Stress

Compared to the vast research on stress at work, research on recovery from work demands is still rare. Often, the research just implicitly assumes that being away from the workplace (such as during vacation, weekends, or during the evening) leads to

recovery from work demands. When an individual is no longer confronted with work demands or work-related demands, recovery can occur. Thus, during leisure time individuals should avoid any work-related demands in order to recover from work demands. We review next empirical evidence on recovery from work demands focusing on recovery during vacations and weekends as well as on recovery during the evening of a workday.

Vacations and Weekends: Effects on Well-Being and Performance-Related Aspects

Vacations as a time away from work give the opportunity for people to recover from work demands and to build new resources. Empirical research on different employee groups and from different countries indicates that vacations are associated with an increase in employee health and well-being. For example, measuring levels of burnout before, during, and after vacation, a study in Israel found a decrease in self-reported burnout after vacation. Furthermore, vacations seem to have a positive impact on psychophysiological indicators such as the immune system or even mortality.

So far, there is only little research on effects of vacation on performance-related outcomes. One study examined the impact of vacations on individual perceptions of effort and task performance at work. Comparison between pre- and postvacation measures indicated no change in task performance. However, study participants reported needing much less effort to get their work done after vacation compared to before vacation. A study measuring changes in employee absenteeism rates found a decrease in absenteeism after vacation compared to prevacation. In summary, although there are the first indications that vacations have a positive effect on performance-related variables, more research seems necessary.

Research further suggests that the positive effects of a vacation subside within a short while after the employee's return to work. This means that well-being and performance-related outcomes return to their prevacation or chronic level. Such fade-out effects have been found within the first few weeks after vacation. Future research should include additional measurement occasions after vacation to better understand the fade-out of recovery effects over time and to possibly reveal factors that may impair or accelerate the fade-out process.

In addition to research on the general effects of vacation, there is some research examining specific variables in the context of a vacation that may promote recovery from work demands. For example, a study found that thinking positively about one's job, pursuing activities that include the experience of mastery, and relaxation during vacation were associated with higher levels of individual well-being after vacation. In contrast, reflecting about one's job in a negative way and facing high levels of nonwork hassles during vacation, as well as high levels of workload directly after vacation, were detrimental for employee well-being and for performance-related variables directly after vacation. In addition, research revealed a positive relationship between people's satisfaction with the vacation and job and their satisfaction with life after vacation. However, more research is necessary to more clearly understand the processes during vacation that underlie changes in well-being and performance-related outcomes.

So far, research on recovery during the weekend is rare. However, one study found effects of specific experiences during the weekend on well-being and performance-related variables after the weekend. Specifically, reflecting positively about one's job and social activities seemed to promote recovery during the weekend, whereas nonwork hassles during the weekend seemed to be detrimental for recovery from work demands. There clearly is the need for more research on recovery from work demands during weekend to better understand the processes that enhance or hinder recovery processes and subsequent well-being and performance.

Recovery during the Evening of a Workday: Effects on Well-Being and Performance-Related Aspects

Recently, research has started examining how employees unwind from work during the evening of a workday. Research indicates positive effects of recovery during the evening on feelings of well-being at bedtime. Specifically, low-effort activities and social and physical activities seem to promote recovery and increase subsequent well-being. In contrast, household and work-related activities seem to impair recovery from work demands, as indicated by lower levels of well-being at bedtime. In addition, generally experiencing off-work activities as positive or as recovering seems to positively impact well-being at bedtime.

With regard to performance-related aspects, a study demonstrated that higher levels of individual regeneration in the morning due to recovery activities the evening before were positively associated with work engagement and proactive behaviors during the workday. Future research should examine the processes underlying those positive effects in more detail.

There is some empirical evidence that the choice of recovery activities during the evening and their effectiveness may be affected by work conditions. For example, job stressors (e.g., time pressure and low

control) seem to be associated with higher levels of job-related rumination after work, more difficulties detaching from work, and less time spent on sport activities. In addition, long work hours seem to be related to a lower capability to relax after work and again to more difficulties detaching from work and less time spent on sport activities.

Finally, the choice of recovery activities and their effectiveness may be affected by interindividual differences. For example, employees who are high in job involvement seem to detach less from work in the evening than employees low in job involvement. Research further indicates a positive relationship between the individual expectation to benefit from the recovery activities and detachment from work during the evening.

Conclusion

Research on stress recovery is a highly important topic that addresses both fundamental and applied questions. With respect to more fundamental research, theoretical models and empirical studies on recovery will add to our full understanding of how stressors affect the organism. Increasing empirical evidence suggests that not just the immediate reaction to a stressor, but processes related to stress recovery play a crucial role in the development of diseases. With respect to more applied questions, it is important to study in more detail the factors that promote quick recovery. Here, a combination of laboratory studies and studies in real-world field setting seem to be particularly promising.

Further Reading

Burke, H. M., Davis, M. C., Otte, C., et al. (2005). Depression and cortisol responses to psychological stress: a meta-analysis. *Psychoneuroendocrinology* **30**, 846–856.

Dickerson, S. S. and Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin* **130**, 355–391.

Eden, D. (2001). Vacations and other respites: studying stress on and off the job. In: Cooper, C. L. & Robertson, I. T. (eds.) *International review of industrial and organizational psychology*, pp. 121–146. Chichester, UK: Wiley.

Fritz, C. and Sonnentag, S. (in press). Recovery, well-being, and performance-related outcomes: the role of workload and vacation experiences. *Journal of Applied Psychology.*

Ganster, D. C., Fox, M. L. and Dwyer, D. J. (2001). Explaining employees' health care costs: a prospective examination of stressful job demands, personal control, and physiological reactivity. *Journal of Applied Psychology* **86**, 954–964.

Glynn, L. M., Christenfeld, M. and Gerin, W. (2002). The role of rumination in recovery from reactivity: cardiovascular consequences of emotional states. *Psychosomatic Medicine* **64**, 714–726.

Gump, B. B. and Matthews, K. A. (1999). Do background stressors influence reactivity to and recovery from acute stressors? *Journal of Applied Social Psychology* **29**, 469–494.

Hocking Schuler, J. L. and O'Brian, W. H. (1997). Cardiovascular recovery from stress and hypertension risk factors: a meta-analytic review. *Psychophysiology* **34**, 649–659.

Linden, W., Earle, T. L., Gerin, W., et al. (1997). Physiological stress reactivity and recovery: conceptual siblings separated at birth? *Journal of Psychosomatic Research* **42**, 117–135.

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* **338**, 171–179.

Pieper, S. and Brosschot, J. F. (2005). Prolonged stress-related cardiovascular activation: is there any? *Annals of Behavioral Medicine* **30**, 91–103.

Sluiter, J. K., Frings-Dresen, M. H. W., Meijman, T. F., et al. (2000). Reactivity and recovery from different types of work measured by catecholamines and cortisol: a systematic literature overview. *Occupational and Environmental Medicine* **57**, 298–315.

Sonnentag, S. (2003). Recovery, work engagement, and proactive behavior: a new look at the interface between non-work and work. *Journal of Applied Psychology* **88**, 518–528.

Stewart, J. C., Janicki, D. L. and Kamarck, T. W. (2006). Cardiovascular reactivity to and recovery from psychological challenge as predictors of 3-year change in blood pressure. *Health Psychology* **25**, 111–118.

Westman, M. and Eden, D. (1997). Effects of a respite from work on burnout: vacation relief and fade-out. *Journal of Applied Psychology* **82**, 516–527.

Reductive Stress

J P Kehrer
Washington State University, Pullman, WA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J P Kehrer, volume 3, pp 327–332, © 2000, Elsevier Inc.

Introduction
Reduction–Oxidation Chemistry
Definitions
Assessing Reductive Stress
Causes of Reductive Stress
Effects of Reductive Stress
Energy and Cell Damage
Pyridine Nucleotides and Cell Damage
Regulation of Gene Expression
Conclusion

Glossary

Free radicals	Molecules with one or more unpaired electrons that are not controlled by their immediate environment.
Oxidative stress	A disturbance in the prooxidant-antioxidant balance in favor of the former that enhances the potential for damage.
Reactive oxygen species	Higher-energy-state forms of oxygen or partially reduced forms of molecular oxygen that have increased reactivity.
Redox balance	The balance between oxidation and reduction states of cellular molecules. Because some molecules must be oxidized when another is reduced, the overall redox balance is inviolable. However, this balance can be altered between subcellular locations or cells.
Reductive stress	A disturbance in the prooxidant-antioxidant balance in favor of the latter that enhances the potential for damage; the result of a cell (or subcellular location) becoming relatively more reduced than normal.

Introduction

The ultimate goal of understanding the biochemical and molecular changes associated with stress is to find pharmacological or gene-based methods of protecting at risk tissues from injury or reversing any injury once it has occurred. The induction of stress at the cellular level can occur as the consequence of exposure to various xenobiotics, physical agents such as heat or radiation, or certain diseases. The specific focus of the current chapter is reductive stress. This type of stress involves alterations in how and where reducing equivalents are generated and used by tissues as cofactors in normal metabolic processes, in various defense systems (particularly oxidant defenses), and in controlling various signal transduction and gene regulation pathways.

Because reducing equivalents are intimately involved in mitochondrial respiration and the generation of energy, reductive stress is most commonly generated through alterations in these pathways, including simple oxygen deprivation. This later process is often followed by reoxygenation, which can produce a sudden burst of oxidants. Thus, reductive stress and oxidative stress are inextricably linked.

Reduction–Oxidation Chemistry

The chemistry of oxidation–reduction reactions is straightforward and easily defined. Specifically, reduction is defined as the addition of one or more electrons to a molecule, whereas oxidation is defined as the removal of one or more electrons. Basic chemical principles dictate that a balance between oxidation and reduction is required in all chemical reactions. Thus, when one substance is oxidized, some other must be reduced.

At the cellular level, life can be defined as the ability to maintain gradients across membranes. Such gradients require an input of energy and may involve ions or molecules. Although at a global level the maintenance of an overall redox balance is inviolable, such a balance is a steady state and can be shifted among different classes of molecules that have different impacts on cell function or between cellular compartments. As a result, individual cells, or areas within a cell, can experience reductant (and obviously concomitant oxidant) imbalances. Significantly, these imbalances do not necessarily deviate from neutrality because, under normal conditions, the presence of substrates used by cells to generate energy means that the intracellular environment of a cell is reduced relative to that found extracellularly.

Because of the reducing nature of the intracellular environment, imbalances toward oxidation tend to be both more common and more studied. A search of the Medline database through March 2006 turned up 38 902 papers using the keyword term oxidative stress but only 43 using the term reductive stress. Despite this huge imbalance in emphasis, the reducing side of the equation is important when it comes to stress and the molecular regulation mechanisms of gene expression.

Definitions

The term oxidative stress was initially defined, by Helmut Sies in 1985, in biological systems as "a disturbance in the prooxidant-antioxidant balance in favor of the former." This definition was later extended to include the adverse functional consequences of such stress by adding "leading to potential damage."

The first use of the term reductive stress was by Albrecht Wendel in 1987. Interestingly, the term has never been precisely defined, perhaps because of the obvious inverse relationship to oxidative stress or perhaps because reductive stress has not been widely considered experimentally or theoretically. In general, reductive stress has been related to conditions of respiratory inhibition and refers to situations in which the status of a cell (or subcellular compartment) becomes relatively more reduced. Such situations actually end up enhancing the formation of reactive oxygen species (ROS), making the links between reductive and oxidative stress even stronger.

There is a strong tendency to imply that stress and damage are synonymous. Overcoming this tendency is difficult because damage clearly is accompanied, and usually preceded, by some sort of stress. A definition of damage may help, but there is no consensus as to what constitutes damage in a biological system. The definition of oxidative (and by analogy, reductive) stress given here indicates that stress only enhances the potential for damage, not that damage has occurred. Because the tissue changes that are routinely measured to assess reductive stress have been also observed in situations that are not necessarily detrimental to tissue or cell functions, it is obvious that stress and damage are not synonymous.

Injury, damage, toxicity, and related terms are qualitative descriptors of changes in various indices that are measured to assess cell or tissue function. Just how significant any particular change is in terms of function is rarely assessed, and thus their meaning is a value judgment made by an individual investigator. It is also important to differentiate adaptive responses from acute responses to a stress. Whether or not any observed changes are directly deleterious to function is also important, as is the time frame being considered because the persistence of some nitially nondamaging adaptation can become detrimental. One useful working definition of injury is any alteration that diminishes essential functions, whether or not cell death ensues. Of course, it must be remembered that even cell death is not always undesirable, as evidenced by apoptosis under normal conditions.

Assessing Reductive Stress

A measurable shift to a more reduced steady state of some biomolecular pool is a reasonable working definition of reductive stress. The biomolecule selected for such a determination will probably differ from that used to identify the presence of oxidative stress. For example, glutathione disulfide (GSSG) is often measured as an index of oxidative stress because it only forms as the result of oxidative processes. Measuring increases in reduced glutathione (GSH) as an index of reductive stress is not, however, feasible because of cellular controls on GSH content and the large basal levels of this thiol (in contrast to GSSG). This means that small changes would need to be detected on a large background. In general, assessing the relative NAD(P)(H) pools has proven to be the most widely used indicator of reductive stress. Other indicators include an increase in the lactate/pyruvate ratio or the β-hydroxybutryate/acetoacetate ratio.

Causes of Reductive Stress

Oxygen Deprivation

The proximate cause of reductive stress during oxygen deprivation (e.g., hypoxia or ischemia) is the loss of the terminal electron acceptor (oxygen) and the resultant accumulation of reduced cofactors (such as NADH and NADPH) along with oxidized substrates. Because this process occurs in the mitochondria, it is this organelle that first exhibits measurable changes, mainly the accumulation of the pyridine nucleotide, NADH. In the whole cell, there is also a loss of high-energy phosphates (mainly ATP). Glycolysis can modify the loss of ATP (although not enough to maintain long-term tissue viability) but at the expense of generating a further quantity of reducing equivalents. Because different tissues, as well as tumor cells versus normal cells, have differing glycolytic capacities, the extent of reductive stress induced by oxygen deprivation varies. There is no evidence to suggest that the oxidized substrates or reduced cofactors are directly toxic, but it is possible that excess reducing equivalents in the form of NADH might drive reactions in directions not normally seen, thereby affecting cell function and ultimately viability. Mechanisms of such effects could involve the generation of free radicals or the disruption of redox-regulated signal transduction pathways (**Figure 1**).

The ability of a tissue to maintain a normal redox balance is believed to be energy-dependent. This concept is supported by studies showing oxidants are more toxic to energy-depleted hypoxic cells than to

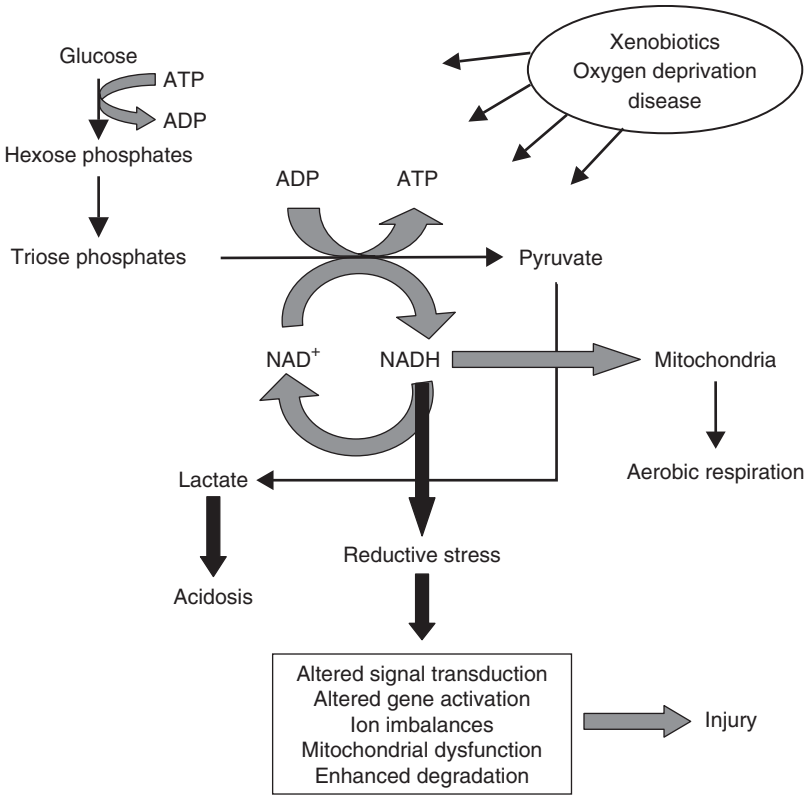


Figure 1 Reductive stress is a situation where the status of a cell (or subcellular compartment) becomes relatively more reduced. Causes of reductive stress can include hypoxia or ischemia and xenobiotics that affect pathways creating and/or using reducing equivalents. As the intracellular milieu becomes more reduced, there can be changes in various signaling pathways that required a defined redox state to be active. Although such effects may be detrimental, the presence of extra reducing equivalents may enhance a cell's ability to respond to an oxidative stress.

normal cells. However, although the production of energy and reducing equivalents are closely linked, they are separate biochemical events and alterations in the supply of reducing equivalents alone has the potential to induce stress.

Xenobiotics

Reductive stress in response to xenobiotics closely resembles that associated with oxygen deprivation. In fact, a widely studied model of reductive stress is termed chemical hypoxia and involves the blockade of mitochondrial respiration at several sites using cyanide and iodoacetate. Although differing from hypoxic hypoxia in some respects (e.g., chemical hypoxia appears to have more effects on NADPH than does oxygen deprivation), the overall biochemistry appears to be similar. Blocking specific mitochondrial electron transport sites with agents such as rotenone or antimycin A may also generate a reductive stress, although it is likely to have some differing effects based on the unique features of these compounds. In fact, such blockage in an aerobic environment may actually increase the generation of ROS.

Effects of Reductive Stress

The gross adverse consequence associated with prolonged oxygen deprivation is obvious – cell or tissue death. However, the biochemical and molecular changes that mediate this gross effect are much less clear. Changes consistent with reductive as well as oxidative stress have been found during hypoxia in whole animals, heart tissue, lung tissue, and chemically hypoxic hepatocytes. It also appears that disrupting cellular energy supplies, and thus reducing equivalent levels, has effects on other redox systems within the cell. For example, mitochondrial thiols are more susceptible to oxidation than are those in the rest of the cell.

Oxidative changes, or an increased susceptibility to oxidation, are consistent findings in various models of reductive stress. The oxidative changes observed in systems in which oxygen metabolism is inhibited could arise from the increased production of ROS or by modifying the cells' ability to respond to other stresses, leading to alterations in redox balance systems. The rate of mitochondrial ROS generation increases with the extent of reduction of mitochondrial

components. Although the low K_m of cytochrome oxidase for oxygen should ensure the removal of virtually all of this molecule during oxygen deprivation, it is possible that the increasingly reduced respiratory chain reacts with residual oxygen to generate ROS. It is also possible that the observed oxidative changes are a consequence of ROS generated on the resumption of respiration. Endogenous antioxidant systems may be impaired, but may continue to function during ischemia. Changes in the content of pyridine nucleotides alone appears not to lead to irreversible injury, but an altered ability to supply reducing equivalents in response to various demands may be critical.

In general, researchers have concluded that oxygen deprivation injury results not only from ATP depletion but also from reductive stress and the consequent generation of ROS by mitochondria. The site of this generation appears to be mainly at complex III, although other cellular sources of ROS appear to exist. It should be noted that not all studies support an increased mitochondrial production of ROS. For example, ROS production may not be increased following hypoxia or ischemia in the heart, brain, or isolated cardiac mitochondria.

Although clearly abnormal and likely to be detrimental in the long term, it is possible that the increased state of reduction that occurs subsequent to oxygen deprivation is actually protective against an acute oxidative stress. Similarly, under conditions of reductive stress, oxidizing agents could, in part, be beneficial because they would provide a sink for excess reducing equivalents.

The detrimental effects of reductive stress may partially be mediated by secondary effects that enhance the generation of ROS. Specifically, reductive stress can result in the intracellular release of iron which, in turn, enhances the production of ROS. Although the mechanisms are not fully understood, excessive NADH generated during a reductive stress can release ferrous iron from ferritin and can also reduce chelated ferric iron. The presence of increased amounts of free iron can, in turn, enhance the transfer of electrons to inappropriate sites, particularly in the presence of oxygen, leading to damage.

Energy and Cell Damage

Damaging events that affect cellular energy levels include mitochondrial dysfunction and the inhibition of glycolysis. Any such event will clearly affect how cells handle their supply of reducing equivalents, which is critical for cellular energy production. As cells adjust to compensate for a damaging event, reducing equivalents are shunted to functions other than producing energy. Such shifts compound the loss of energy and must be considered potentially damaging events. Cell injury processes targeted to energy production may, therefore, represent a unique challenge for cells to increase their supply of reducing equivalents while maintaining critical redox balance to avoid further disruption of function.

The role of energy in cell injury has been studied in a variety of systems. The importance of secondary messengers in mediating such injury is well recognized. Of significance is calcium, whose intracellular levels appear to be altered in response to redox reactions as well as in the presence of energy deficiency. Changes in the levels of this cation are important for various signal transduction pathways. Thus, the regulation of signal transduction pathways by reductants is not necessarily direct, but may be mediated through secondary messengers such as calcium signaling and protein phosphorylation. Cellular pH is also important because an acidotic environment is protective against hypoxic or ischemic injury.

Pyridine Nucleotides and Cell Damage

The maintenance of reduced intracellular pyridine nucleotides appears to be a key factor in preserving cell viability because chemicals such as menadione, paraquat, and hydrogen peroxide deplete them before the loss of ATP and cytotoxicity. The key pathways maintained by NAD(P)(H) are not known, although modification of NAD(P) reduction status may lead to changes in the redox state of mitochondrial thiol groups. This, in turn, may affect cellular free calcium levels, the mitochondrial permeability transition, and trigger apoptosis.

Regulation of Gene Expression

Genes

The activation of gene expression in response to stress appears to be a universally conserved response in cells and substantial data suggest that oxidants can mediate gene activation. Whether a similar group of reductive stress-responsive genes exists is not known. It has been proposed that *grp* (glucose-regulated protein) genes are more responsive to reductive stress, whereas *hsp* (heat shock protein) genes respond better to oxidative stress.

The role of thiols in the control of cell growth and death is well accepted, although the specific pathways involved are not yet clear. Glutathione, thioredoxin reductase/thioredoxin, protein disulfide isomerase, glutaredoxin, and redox factor-1 have all been shown to affect protein thiols, thereby potentially

regulating protein folding, assembly, activity, and binding as transcription factors to specific DNA sequences. Although most work has focused on changes in terms of oxidants, it is again obvious that the reducing side of the equation could also be important. This is particularly true with thioredoxin, whose roles in regulating cell signaling are becoming increasingly appreciated.

Transcription Factors

At least 64 redox-regulated genes and transcription factors have been identified. More so than most other transcription factors, there is an extensive base of information regarding the activation of NF-κB. This activation requires phosphorylation of the inhibitory IκB subunit, which results in its dissociation from the inactive complex and its degradation by proteasome. A role for oxidants in NF-κB activation is based on the observations that oxidizing conditions activate NF-κB in several cell types, antioxidants can block this activation, and oxidant production is enhanced by various NF-κB inducers such as tumor necrosis factor α. However, at the nuclear level, NF-κB must be reduced in order to bind to DNA. Thus, the overall redox status of specific subcellular sites is crucial for determining the activation state of NF-κB. Compartmentation may also be critical for the redox responses of NE-F2-related factor 2 (Nrf-2). It has been suggested that Nrf-2 is retained in the cytosol by Keap-1 but that, when critical thiols on Keap-1 are oxidized, Nrf-2 is released and translocated to the nucleus, where Nrf-2 participates in the transcriptional activation of numerous genes. For Nrf-2 to be fully functional in transcriptional activation, a key cysteine residue must be in the thiol (reduced) form. Although phosphorylation may also be involved, this model further highlights the importance of reductive reactions in controlling signal transduction.

Activator protein 1 (AP-1) is a ubiquitous collection of protein complexes known to regulate transcription in response to environmental stimuli. It is composed of various gene products from the *fos* and *jun* proto-oncogene families. The products of these genes form homodimeric (Jun–Jun) and heterodimeric (Fos–Jun) complexes that bind to DNA, suggesting a crucial role of AP-1 in the control of cell proliferation. In addition to being stimulated by a wide range of xenobiotics or factors that promote cell proliferation, the hypothesis that cellular redox state plays an important role in the activation of AP-1 is well supported. AP-1 appears to be part of a general mechanism of regulation of gene expression following an alteration in redox balance. However, the pathways that control this expression are not clear. Antioxidants strongly activate AP-1, yet growth factor-induced AP-1 activation appears to be ROS-dependent. Thus, although AP-1 and NF-κB seem to respond oppositely to antioxidants, reductants and oxidants may have a role in the activation process of both factors.

GSH is clearly important for many of the reduction reactions required for normal cell functioning. In addition, recent data have demonstrated that the thioredoxin system may be the more proximal regulatory factor for several transcription factors including NF-κB. The overall reduction status of a cell and, in particular, the relative state of both the cytosol and nucleus in terms of thiol redox balance affect the activity of transcription factors. Reductive stress is, therefore, expected to have an effect on their activities.

Conclusion

Reductive and oxidative stress are closely intertwined processes that can affect cells at various levels. Although the oxidative side of the equation has been the most studied, the reductive side has the significant potential to regulate signal transduction pathways and gene expression, as well as more obvious cellular stress responses.

See Also the Following Articles

Apoptosis; Oxidative Stress.

Further Reading

Allen, R. G. and Tresini, M. (2000). Oxidative stress and gene regulation. *Free Radical Biology and Medicine* **28**, 463–499.

Arrigo, A-P. (1998). Small stress proteins: chaperones that act as regulators of intracellular redox state and programmed cell death. *Biological Chemistry* **379**, 19–26.

Cross, J. V. and Templeton, D. J. (2004). Thiol oxidation of cell signaling proteins: controlling apoptotic equilibrium. *Journal of Cellular Biochemistry* **93**, 104–111.

Dalton, T. P., Shertzer, H. G. and Puga, A. (1999). Regulation of gene expression by reactive oxygen. *Annual Review of Pharmacology and Toxicology* **39**, 67–101.

Dawson, T. L., Gores, G. J., Nieminen, A-L., et al. (1993). Mitochondria as a source of reactive oxygen species during reductive stress in rat hepatocytes. *American Journal of Physiology* **264**, C961–C967.

Finkel, T. (2003). Oxidant signals and oxidative stress. *Current Opinion in Cell Biology* **15**, 247–254.

Forman, H. J., Torres, M. and Fukuto, J. (2002). Redox signaling. *Molecular and Cellular Biochemistry* **234–235**, 49–62.

Haddad, J. J. (2002). Antioxidant and prooxidant mechanisms in the regulation of redox(y)-sensitive transcription factors. *Cell Signaling* **14**, 879–897.

Halleck, M. M., Holbrook, N. J., Skinner, J., et al. (1997). The molecular response to reductive stress in LLC-PK1 renal epithelial cells: coordinate transcriptional

regulation of gadd153 and grp78 genes by thiols. *Cell Stress & Chaperones* **2**, 31–40.
Jones, D. P. (1997). New concepts of the molecular pathenogenesis arising from hypoxia. In: King, T. E., Mason, H. S. & Morrison, M. (eds.) *Oxidases and related redox systems*, pp. 127–144. New York: A. R. Liss.
Kehrer, J. P. and Lund, L. G. (1994). Cellular reducing equivalents and oxidative stress. *Free Radical Biology and Medicine* **17**, 65–75.
Sies, H. (ed.) (1985). *Oxidative Stress*. p. 1. New York: Academic Press.
Turrens, J. F. (2003). Mitochondrial formation of reactive oxygen species. *Journal of Physiology* **522**(2), 335–344.
Watson, W. H., Yang, X., Choi, Y. E., et al. (2004). Thioredoxin and its role in toxicology. *Toxicological Sciences* **78**, 3–14.

Reenactment Techniques

N Wong
Uniformed Services University of the Health Sciences, Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by N Wong, volume 3, pp 333–334, © 2000, Elsevier Inc.

Rationale for Reenactment
Common Techniques

Glossary

Abreaction	A process by which repressed material, particularly a painful experience or conflict, is brought back into consciousness. In the process the person not only recalls, but also relives the situation and experiences the appropriate emotional response.
Catharsis	The verbalization of ideas, thoughts, and expressed material that is accompanied by an emotional response and that produces a state of relief to the individual.
Psychodrama	A form of group therapy devised by Moreno in the 1920s drawing upon theatrical techniques to explore a patient's personality makeup, interpersonal relationships, conflicts, and emotional patterns.

Rationale for Reenactment

When exposed acutely or chronically to stressful or traumatic situations, some individuals may develop physical and emotional symptoms, which may range from bothersome to totally incapacitating. These symptoms may happen occasionally, or they can be present throughout the day and night. The aim of having patients reenact their trauma is to prevent, diminish, or totally alleviate any potential or actual symptoms and emotional suffering. In addition to the use of psychodynamic therapy, several categories of reenactment techniques exist and are applied with respect to the clinical indication and ego strength of the individual.

Common Techniques

Debriefing

One of the most commonly utilized techniques to help individuals reenact or recall an acute or intermediate traumatic situation, debriefing can be applied in an individual or group setting. Participants are encouraged to retell their personal experiences and thereby gain an emotional release, or catharsis, that hopefully prevents or reduces the risk for more serious stress reactions. In the process, there is a sharing and education of those experiencing a common trauma that promote ventilation of feelings through the expression of emotions and the response of others.

Psychodrama

Psychodrama is a form of group therapy, but some of its techniques, such as warming up, scene setting, and therapeutic soliloquy, can be used in the individual setting as well for patients suffering from stress. Other psychodrama techniques are role playing, role reversal, therapeutic soliloquy, and mirror. The therapist or leader is an active participant. In role playing, the patient is asked to perform an impromptu dramatization of the traumatic situation. In role reversal, another individual in the group attempts to reenact what the subject may be thinking, or feeling, in the presence of the subject. In therapeutic soliloquy, the individual is encouraged to verbalize side dialogues and actions during a psychodrama scene. The mirror

technique consists of having another person reenact what the subject has done, presenting identical behaviors and expressing similar feelings to show how others experience the patient while the subject sits behind a mirror.

Hypnosis

This modality is especially useful in the reenactment of stressful or traumatic situations causing dissociative amnesia and dissociative fugue with memory loss. Hypnosis is best utilized to explore the events that precipitated the fugue or amnesia. When employing this technique, clinicians should suggest to the hypnotized patients that they may choose to forget some or even all of what was remembered or reenacted in the trance state to prevent them from being overwhelmed by the emergence of unconscious material that the patients are not yet ready to handle. The recaptured or selectively shared data from patients are then dealt with in the fully conscious state to help them cope with the underlying conflicts. It is generally advisable to employ other conscious methods before resorting to the use of hypnosis or chemical means to bring about a reenactment of a traumatic situation. While hypnosis is usually safe to use with most healthy individuals, the therapist or treater should exercise the same degree of caution in probing unconscious material in psychotherapy.

Medication-Facilitated Interviews

Clinicians may use intravenous drug-induced interviews to help patients recall or reenact traumatic situations or conflicts when symptoms may be disabling, or when great resistance is encountered in examining or recalling memories. Commonly utilized drugs include amobarbital, pentobarbital, intravenous lorazepam, and hexobarbital. A disadvantage of these interviews lies in the chemically induced amnesia that often occurs. Also, because of the risk of adverse reactions to the drugs, medication-facilitated interviews should only be performed by trained individuals with immediate availability of life support systems and the presence of an anesthetist.

Use of Media

Young children may not have the verbal capacity to fully express their anxieties and emotional suffering when exposed to traumatic or stressful situations. Therapy may be conducted via play, painting and drawing, participation in story telling, creating illustrated storybooks, and role playing through which the trauma may be reenacted, explored and clarified, and resolved. For older children, adolescents, and adults, the creation of a documentary video or a play reenacting the trauma can prove extremely therapeutic, providing opportunities for ventilation, abreaction, clarification, reality testing, and mastery.

See Also the Following Articles

Hypnosis; Trauma and Memory.

Further Reading

Freehill, K. M. (1992). Critical incident stress debriefing in health care. *Critical Care Clinic* **8**, 491–500.

Hanney, L. and Kozlowska, K. (2002). Healing traumatized children: creating illustrated storybooks in family therapy. *Family Process* **41**, 37–65.

Krietemeyer, B. C. and Heiney, S. P. (1992). Storytelling as a therapeutic technique in a group of school-aged oncology patients. *Child Health Care* **21**, 14–20.

Moreno, J. L. (1947). *Psychodrama.* New York: Beacon Press.

Quittner, A. L., Tolbert, V. E., Regoli, M. J., et al. (1996). Development of the role-play inventory of situations and coping strategies for parents of children with cystic fibrosis. *Journal of Pediatric Psychology* **21**, 209–235.

Rhue, J., Lynn, S. and Kirsch, I. (1988). *Handbook of clinical hypnosis.* Washington, D.C.: American Psychological Association.

Tosone, C., Gelman, C. R. and McVeigh, L. (2005). Through their own eyes: a media-base group approach to adolescent trauma. *International Journal of Group Psychotherapy* **55**, 415–432.

Refugees, Stress in

J D Kinzie
Oregon Health and Sciences University, Portland, OR, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J D Kinzie, volume 3, pp 335–337, © 2000, Elsevier Inc.

Scope of the Problem
Stress in Refugees
Stress-Related Disorders in Refugees
Conclusion

Glossary

Depression	A specific psychiatric disorder consisting of lowered mood and symptoms, such as decreased interest in most activity, loss of appetite and weight, poor sleep, fatigue, feelings of worthlessness, poor concentration, and recurrent thoughts of death.
Posttraumatic stress disorder	A specific psychiatric disorder following psychological trauma with symptoms of re-experiencing, such as nightmares and flashback; avoidance behavior, such as attempts to avoid thoughts and memories of the events; and hyperarousal symptoms, such as irritability, startle response, and difficulty sleeping.
Refugee	A person who flees to another country as a result of, or to avoid, political prosecution, war, starvation, or other disasters.
Schizophrenia	A severe mental disorder characterized by loss of contact with reality, hallucination, delusions, and often impairment of social and vocational functioning.

Technically, a refugee is a person who flees to another country usually to escape political prosecution, war, starvation, and other disaster. This definition does not include those displaced internally in their own country. Refugees are distinguished from migrants who voluntarily migrate to another country for economic education or personal reasons. A special group of refugees are the asylum seekers who flee to a foreign country without official recognition or on a temporary visa status and request refugee status usually because of perceived dangers if they return to their homeland. While their claims are reviewed, asylum seekers live in an ambiguous status with the threat of forced repatriation.

Scope of the Problem

At the end of 2004, the United Nations High Commission of Refugees reported 19.2 million persons of concern (including refugees, asylum seekers, stateless, and other). This was a 13% increase from 2003. Those with formal refugee status comprise 48% of the 19.2 million but many stateless persons have not been systemically identified and therefore the total persons of concern is probably much greater than 19.2 million. The areas of southern Africa and CASWANAME (Central Asia, South West Asia, North Africa and Middle East) each hosted 30% of the global refugee population. The Americas only hosted 7%. Afghanistan has been the largest country of origin of refugees despite 3.5 million Afghans in the past 3 years voluntarily repatriated home. The continual civil wars and ethnic strife in Africa and the Middle East indicates the worldwide problems of refugees and asylum seekers will continue to exist.

Stress in Refugees

Types of Stress

The stresses on refugees are diverse, multiple, and frequently catastrophic in nature. The phases of stress are often divided into preflight; flight and separation; asylum; and resettlement. Preflight stress includes poor economic conditions, food shortage, famine, political persecution, social upheaval, and violence. The results can include hunger, malnutrition, physical disability, and physical and psychological trauma. The flight and separation phase can include separation from family and society, sexual assault, violence, and collapse of social support. The result can be multiple physical and social traumas and the sense of being isolated and alone. The asylum phase often includes temporary resettlement in a refugee camp and can involve the threats of being forced to return home, inhospitable living conditions, unemployment, food shortages, and lack of health services. The results can include malnutrition, illness, and despair. Resettlement stress, often in a third country, can include unemployment, social isolation, acculturation problems, limited social ties, prejudice, language barriers, intergenerational conflicts, and minority status. The results can be severe economic hardships, social isolation, delinquency among adolescents, generational, and conflicts.

The amount of direct physical and psychological trauma varies among the refugee groups and the

amount of violence prior to and during their escape. Some, such as the first group of Vietnamese refugees, left because of fear of harm rather than direct harm or threats. Others, for example Cambodian refugees, suffered severe stress over 4 years, which involved a total collapse of their society with starvation, torture, witnessing executions, wanton deaths, forced labor for 4 years, and even violence during the escape and refugee camp experience.

There have been mass murders in the name of ethnic cleansing in Bosnia, and murder, starvation, and random lawlessness with gang violence in Somalia. Refugees fleeing severe ethnic wars, and sadistic acts of cruelty in Africa are often displayed on television. Although technically not refugees (they still stay in their own country), survivors of powerful tsunamis, hurricanes, and earthquakes experience many of the same stress as formal refugees. In general victims of acts of God have less psychiatric disturbance than those victims of man-made disasters, i.e., war.

Vulnerable Groups

At different times, different groups can be more vulnerable to stress or persecution. During armed conflict, young men are often singled out as potential enemies and severely persecuted or executed. Women are particularly vulnerable to sexual abuse (rape) when fleeing from their own country or even in settlement camps. At the country of resettlement, women and the elderly are often isolated at home and experience more difficulty in adjusting. Children acculturate faster as a result of school experience, which can greatly increase the intergenerational conflicts in the family over time.

Protective Factors

Since studies during the bombings of London in World War II, it has been shown that children that stay in the care of their mother, even in dangerous bombing situations, do better than those sent to the safe countryside away from families. An intact family or extended family that can maintain daily routines can provide some protection against the most serious psychological effects. Many studies have shown that having effective social linkages and social support provide a sense of belonging and improved identity, which protects against the stress of social adjustment.

Religious affiliations have been found to be protective in some refugee status, and strong ideological and political commitments have even reduced the effects of torture. The more massive and prolonged the psychological trauma, the less effective are any protective mechanisms and the more consistently symptoms are found.

Asylum seekers arrive in a country on a temporary visa (or even illegally) and request refugee status. Usually, during a prolonged waiting period, their claims are reviewed, basic services are denied, and there is even the threat of forced repatriation. This group, which has been studied in several countries, has high postmigration stress and often severe premigration trauma. With an unclear legal status, this group is very vulnerable and may develop an increased rate of psychiatric symptoms.

Stress-Related Disorders in Refugees

Physical Disorders

The first studies of refugees occurred with Nazi concentration camp survivors who fled to another country after World War II. Their multiple symptoms were thought to be related to brain disease due to head trauma or malnutrition. Later, the symptoms were characterized as concentration camp syndrome (which is similar to posttraumatic stress disorder (PTSD)) and were found primarily to be due to the massive psychological trauma of the holocaust. Nevertheless, many refugees do suffer from head trauma, which may explain some of the complaints of memory and concentration problems.

Although long-term health effects are inconsistently found, many refugees report marked somatic distress, and 15–20% report health impairment. Among various groups of refugees (Cambodians, Vietnamese, Russians, Bosnian, and Somalis) hypertension was found in 40–50% and diabetes (type II) was found in 13–17%. Although dietary changes, refugee status, and effects of aging play a role, severe trauma contributes to high prevalence of these two diseases.

Behavioral Disorders

Refugee populations may be particularly sensitive to alcoholism, drug abuse, delinquency, or antisocial behavior. All the losses, including status and employment, may cause adult males to use alcohol and drugs as a temporary escape in both the camps and in the resettlement country. Asian refugees historically have tended to have lower rates of alcoholism, whereas Central American males tend to have high rates of substance abuse as refugees. Young people in resettlement countries have been found to engage increasingly in antisocial behavior and drug abuse. Refugee children can experience a severe disruption of family relationships and much intergenerational conflict in

their new country. A lack of effective role models, a sense of social isolation, and poor academic performance may turn adolescents to use of drugs and drug dealing in an attempt to acculturate to the host society. Such behavior by adult and juvenile males can increase family stress greatly and lead to family breakup and domestic violence.

Indeed in refugee children from Central America domestic violence (often secondary to alcoholism) is more likely to be the major trauma than war-related violence.

Psychiatric Disorders

The most persistent effects on refugees have been the consistent finding of a high rate of psychiatric disorders. Earlier reports from European refugees after World War II indicate an increased rate of schizophrenia. Later studies have shown that those refugees from the Nazi concentration camps were found to have suffered from a specific disorder characterized by fatigue, irritability, restlessness, anxiety, depression, and frequent nightmares. This concentration camp syndrome was found to be related to massive psychological trauma. The syndrome was largely forgotten; but with a large influx of refugees in the United States in the late 1970s and 1980s, it was found that many refugees suffered from PTSD. PTSD is found to be similar, if not identical, to the syndrome called concentration camp syndrome. By definition, PTSD is the result of a traumatic event in which an individual experiences or confronts death or serious injury and responds with fear, helplessness, or horror. The symptoms can cause re-experiencing such as nightmares, avoidance behavior, or a numbing experience such as efforts to avoid thoughts of the events and hyperarousal symptoms such as difficulty falling asleep, irritability, poor concentration, and startle response. In a clinic treating refugees from Southeast Asia, a 54% rate of PTSD was found among the Vietnamese refugees. The traumatic experience was related both to the Vietnam War and to the escape processes. The 92% rate of PTSD among Cambodian patients all related to the Pol Pot concentration camps. Most clinics have found PTSD to be similarly high, although the rates for Cambodians in the community (nonpatients) tend to be lower. A great deal of research and clinical experience indicates that traumatized refugees have high rates of PTSD. Information exists on a number of refugees: Chilean, Salvadorian, Ethiopian Jews, Afghan, Somalis and Bosnians, as well as Tamil and Burmese asylum seekers. A recent community survey of Cambodian refugees two decades after resettlement in the United States found that 62% still suffered from PTSD. This is almost twice as high as among Cambodians living in Cambodia indicating that refugee status and/or immigration stress contribute to ongoing symptoms. The highly televised bombing of Twin Towers on 11 September 2001 was very disturbing to refugees and many (particularly Somalis and Bosnians) had markedly increased symptoms after repeated viewing. The symptoms are often quite persistent and can recur after periods of quiescence.

Depression characterized by lowered mood, low energy, low interest, poor sleep, and poor appetite is also very high among refugees. PTSD has generally been found to be related to specific traumatic events, whereas depression relates more to postmigration stress. Depression tends to improve over time more than PTSD symptoms.

Schizophrenia, a severe mental disorder characterized by delusions and hallucinations, which is found to be high in refugees from World War II, continues to be increased in current refugees. However, the number involved is not nearly as large as those having PTSD and/or depression.

Conclusion

It is now clear that refugees suffer from much stress, sometimes of catastrophic proportions. This can result in symptoms of PTSD and depression, which are similar regardless of the culture of the refugees or the environment where the refugees resettle. As the number of refugees worldwide increases, the number of refugees impaired and suffering will continue to increase

See Also the Following Articles

Schizophrenia; Social Support; Survivor Guilt; War Stress in the Former Yugoslavia; War-Related Posttraumatic Stress Disorder, Treatment of.

Further Reading

Boehnlein, J. K. and Kinzie, J. D. (1995). Refugee trauma. *Transcultural Psychiatric Research Review* **32**, 223–252.

Desjarlais, R., Eisenburg, L., Good, B., et al. (eds.) (1995). *World mental health: problems and priorities in low-income countries*. New York: Oxford University Press.

Jaranson, J. J. and Popkin, M. K. (1998). *Caring for victims of torture*. Washington, DC: American Psychiatric Press.

Marsella, A., Bornemann, T. and Ekblad, S. (eds.) (1994). *Amidst peril and pain*. Washington, DC: American Psychiatric Association.

Wilson, J. P. and Drozdek, B. (eds.) (2004). *Broken spirits: the treatment of traumatized asylum seekers, refugees, war and torture victims*. New York: Brunner-Routledge Press.

Regional Blood Flow, Stress Effects

W W Blessing
Centre for Neuroscience, Flinders Medical Centre, Adelaide, SA, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by W W Blessing, volume 3, pp 338–340, © 2000, Elsevier Inc.

Introduction
Exposure to Cold Causes a Sudden Decrease in Cutaneous Blood Flow
When the Individual is Frightened, Cutaneous Blood Flow Falls
Psychotropic Drugs Alter Cutaneous Blood Flow
Cutaneous Blood Flow Decreases During the Febrile Response to an Infectious Agent
Neuronal Pathways in Brain and Spinal Cord Co-ordinating Sympathetic Outflow to the Cutaneous Vascular Bed
Conclusion

Glossary

Clozapine	A drug used to treat schizophrenia, called an atypical antipychotic agent because it has a reduced tendency to cause muscle rigidity in comparison with classical drugs such as haloperidol.
Raphe/parapyramidal region	Neurons in the midline and ventromedial region of the medulla oblongata with direct axonal projections to the spinal cord.
Thermoregulatory cutaneous vascular beds	Regions of the skin (e.g., fingers in humans and tail in rats) where heat is transferred from the blood to the environment.

Introduction

During the course of evolution, natural selection has shaped behavior as well as psychological and physiological processes so that they operate in an integrated fashion, enabling the individual to successfully undertake the daily life activities necessary for survival and reproduction. During everyday life, the regional distribution of the blood flow is patterned by a combination of central command (brain-initiated) and reflex (organ-initiated) control systems, so that flow to a particular bodily organ is appropriate for the particular activity in which the individual is engaged at a particular time.

Broadly defined, an acute stress may be envisaged as the occurrence of an intrusive event (physical, psychological, or physiological) that could threaten survival of the individual or, in a broader sense, survival of the species. Such events initiate adaptive responses. W. B. Cannon appreciated that the behavioral response to stress (the flight-or-fight response) is accompanied by an acute redistribution of the cardiac output, so that blood is directed to the bodily organs involved in coping with the stressful event. This was an extremely important contribution to physiological thinking, a reminder of the concepts so graphically proposed by Darwin in his book *Expression of the emotions in man and animals*.

Cannon, however, believed the patterns of activity in sympathetic nerves supplying the blood vessels to be diffusely directed rather than regionally focused. The central command process, so Cannon thought, results in a unitary activation of all the sympathetic nerves innervating the adrenal gland and the blood vessels. At the time, this view was based on a general theoretical stance, rather than on experimentally demonstrated changes. Cannon was thinking of a highly mobile animal, actively engaged in fight or flight. In Cannon's scheme, the principal emphasis was on increases in blood flow to vascular beds in skeletal muscle. Measurement of regional blood flow or sympathetic nerve activity in a conscious behaving animal is still technically difficult, so that there is still much to be learnt. In a cat confronted with a second aggressive cat, blood flow to skeletal muscle beds appears to depend on how much the animal actually moves and on whether particular muscle groups are involved in the movement. There is no general increase in blood flow to all skeletal muscles.

Cannon was correct in his view that the response to stressful events includes diversion of blood away from the skin. The short-term (over minutes) nutrient requirements of the skin are not crucial, so that the supplying blood can be diverted elsewhere without deleterious effects to the skin itself. The skin is vulnerable to injury precisely because, as the border between the individual and the external environment, it bears the initial brunt of any actual physical attack upon the individual. If blood is diverted from the skin, the danger of major injury-induced hemorrhage is reduced. This article, drawing on experiments conducted in experimental animals (rabbits and rats), outlines some of the environmental stresses which, via the central command process, activate

cutaneous sympathetic neural outflow so that blood flow to the skin is suddenly reduced.

Exposure to Cold Causes a Sudden Decrease in Cutaneous Blood Flow

The volume of blood flowing through thermoregulatory (heat-exchanging) cutaneous blood vessels (ear pinna in rabbits, tail in rats, fingers and toes in humans) is rapidly reduced when the individual is suddenly exposed to cold (a physical stress), so that heat loss from the body is reduced. The brain, via the various sensory inputs, detects the cold environment and particular brain centers activate the sympathetic nerves innervating the blood vessels, thereby constricting the vessels and reducing the amount of blood that enters the heat-exchanging areas of the body. Since the final control neurons within the central nervous system (the sympathetic preganglionic neurons) are located in the thoracic and upper lumbar segments of the spinal cord, there must be neural connections linking forebrain, midbrain and hindbrain with these spinal neurons. The descending pathways are briefly summarized later in this article.

When the Individual is Frightened, Cutaneous Blood Flow Falls

When the individual detects potentially dangerous events in the external environment (physical or psychological) special alerting rhythms occur in the hippocampal electroencephalograph followed, a few seconds later by a sympathetically mediated fall in cutaneous blood flow (Figure 1a), with a return to baseline levels after a minute or so if the danger is transient. In conscious rabbits (especially bred for use in laboratory experiments and thus familiar with humans) apparently minor stimuli (e.g., the laboratory telephone ringing or an unexpected touch of the fur) often cause cutaneous blood flow to fall suddenly to near zero levels, without change in blood flow to the kidneys, with a small decrease in blood flow to the intestines and abdominal viscera, and with no change in blood flow to vascular beds in skeletal muscle (Figure 1b). With relatively minor alerting stimuli there is little or no change in cardiac output. There may be no obvious behavioral change. If the animal is more strongly alerted, either because the stimulus is more threatening or because the animal is already anxious, the cutaneous vasoconstriction is more prolonged. Heat loss via the cutaneous circulation is thereby interrupted for a longer period. Body temperature may increase, a situation referred to as stress-induced hyperthermia.

One way of introducing a more threatening stimulus is to utilize the conditioned fear response. Rats subjected to an electric shock exhibit behavioral change (freezing) and cardiovascular changes. If the electrical shock is delivered soon after the rat has been moved to an unfamiliar cage, similar behavioral

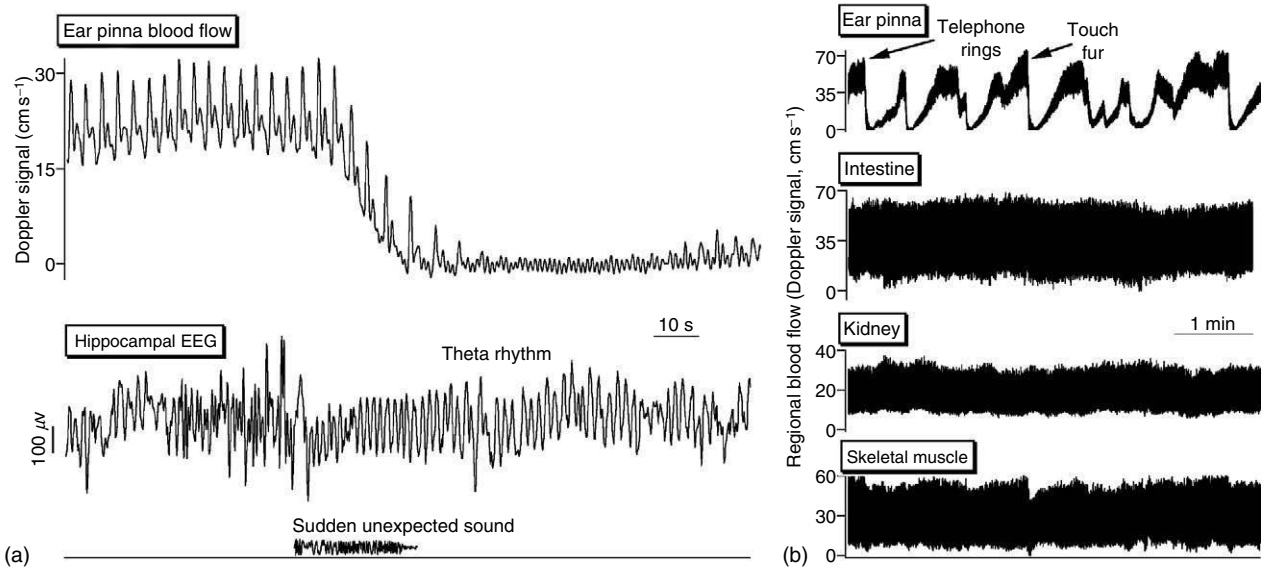


Figure 1 a, Ear pinna arterial blood flow (top panel) and hippocampal electroencephalograph (bottom panel) in a conscious rabbit, recorded before and after onset of a sudden sound. After the sound onset, the electroencephalograph develops a slow (6–7 Hz) regular rhythm indicating that the sound is significant for the rabbit. Approximately 1 s after the sound, ear pinna blood flow falls rapidly to near zero levels. b, Blood flow recorded simultaneously from arteries supplying the ear pinna, intestine, kidney, and skeletal muscle in the hind limb. Ear pinna blood flow falls to near zero levels when the animal detects potentially dangerous environmental stimuli (see Yu and Blessing, 1997: R208).

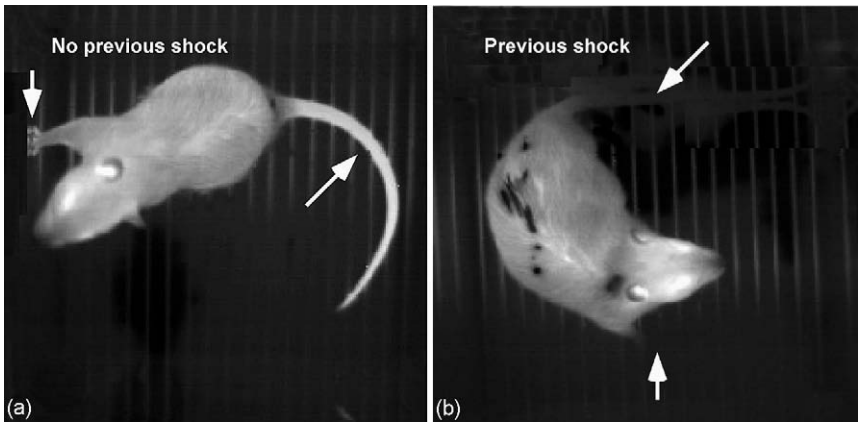


Figure 2 Infrared digital images of one fear-conditioned a, and one sham-conditioned b, rat obtained during re-exposure to the experimental box. Brighter regions indicate increased temperature (i.e., higher blood flow). Note (white arrows) the fall in tail and paw temperature in the fear-conditioned, but not in the sham-conditioned rat during re-exposure to the experimental box (see Vianna and Carrive, 2005: 2505).

and cardiovascular changes occur when the rat is subsequently reintroduced to the same cage, i.e., the rat displays a conditioned fear response. Cutaneous vasoconstriction is a prominent component of the cardiovascular response (Figure 2).

Psychotropic Drugs Alter Cutaneous Blood Flow

There is an intricate interrelationship between neurotransmitters controlling emotional/psychological functioning and neurotransmitters involved in CNS neural pathways regulating cutaneous blood flow. Serotonin (5-hydroxytryptamine, 5-HT) and dopamine are important in both systems. Lysergic acid diethylamide (LSD) and related agents cause hallucinations by an agonist action at subclasses of serotonin receptors (5-HT2A receptors). Ecstasy (methylenedioxymetamphetamine, MDMA) causes release of serotonin in the brain, with subsequent activation of 5-HT2A receptors. These drugs also markedly activate the thermoregulatory sympathetic outflow to the cutaneous vascular bed, causing vasoconstriction and reducing heat loss from the body. Clozapine, a very valuable antipsychotic drug used in the treatment of schizophrenia and schizoaffective disorders to reduce mental stress, reverses the effects of MDMA and LSD-like agents on emotional function. Clozapine also reverses MDMA-induced activation of the cutaneous sympathetic outflow. In drug-free animals, clozapine inhibits resting cutaneous sympathetic outflow, so that cutaneous blood flow increases (Figure 3a), as well as reducing the increase in cutaneous sympathetic discharge normally initiated when the individual detects a frightening stimulus (Figure 3b).

Cutaneous Blood Flow Decreases During the Febrile Response to an Infectious Agent

Infectious agents such as viruses or bacteria cause body temperature to increase (the febrile response). Underlying physiological mechanisms include activation of the cutaneous sympathetic outflow, with constriction of the cutaneous vessels so that body temperature increases partially because heat is no longer transferred from the body to the environment.

Neuronal Pathways in Brain and Spinal Cord Co-ordinating Sympathetic Outflow to the Cutaneous Vascular Bed

Stressful environmental stimuli are detected by the different senses (vision, hearing, touch, temperature, and pain sensitivity, etc.). These afferent stimuli are processed in forebrain centers, including the amygdala and the dorsomedial nucleus of the hypothalamus. Efferent neuronal signals reach the sympathetic preganglionic neurons in the spinal cord by way of a major integrative center in the lower brainstem, in the raphe/parapyramidal region of the rostral medulla oblongata. Neurons in the raphe have direct axonal projections to the sympathetic preganglionic neurons in the spinal cord (Figure 4).

Conclusion

As the individual prepares to cope with the different forms of environmental stress, blood flow to the cutaneous vascular bed is reduced via activation of the relevant sympathetic nerves, a response initiated in the brain and co-ordinated by spinally projecting

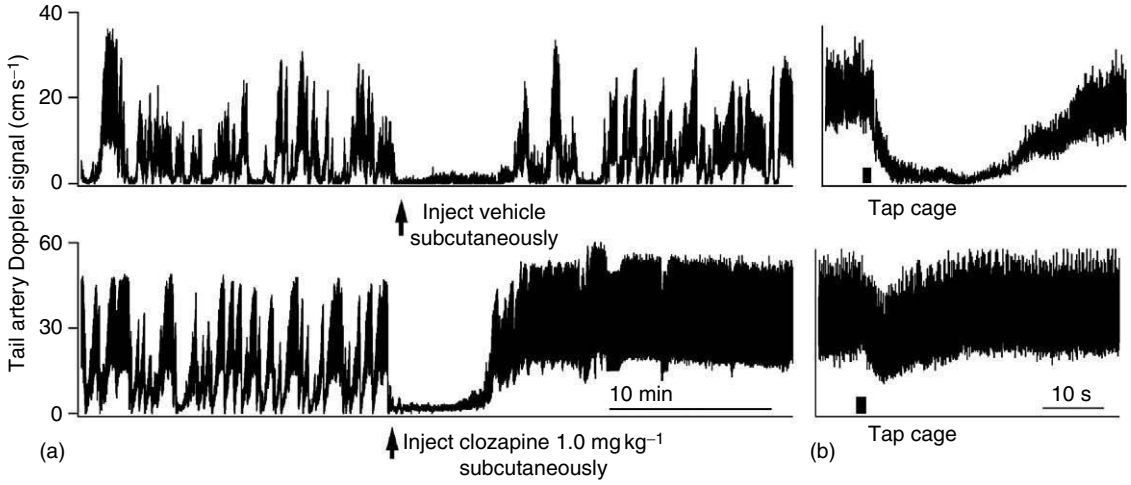


Figure 3 Blood flow to the tail of the conscious freely moving rat measured with a timeframe of minutes (a) and seconds (b). In (a) either vehicle control solution (top panel) or clozapine solution (bottom panel) was administered. Before the injection both traces show frequent sudden falls to near zero levels in response to perception of alarming events by the animal. Tail blood flow falls during the injection procedure in both panels (the rat is alarmed by the procedure). After clozapine, resting blood flow increases and the sudden falls in flow are no longer observed. In (b) the upper panel was obtained from a control, vehicle-treated rat and the lower panel was from a rat treated with clozapine 30 min before. A sudden tap to the animal's cage elicits a rapid fall in tail blood flow, to near zero levels, in the vehicle-treated animal but not in the clozapine-treated animal (see Blessing, 2005: 518).

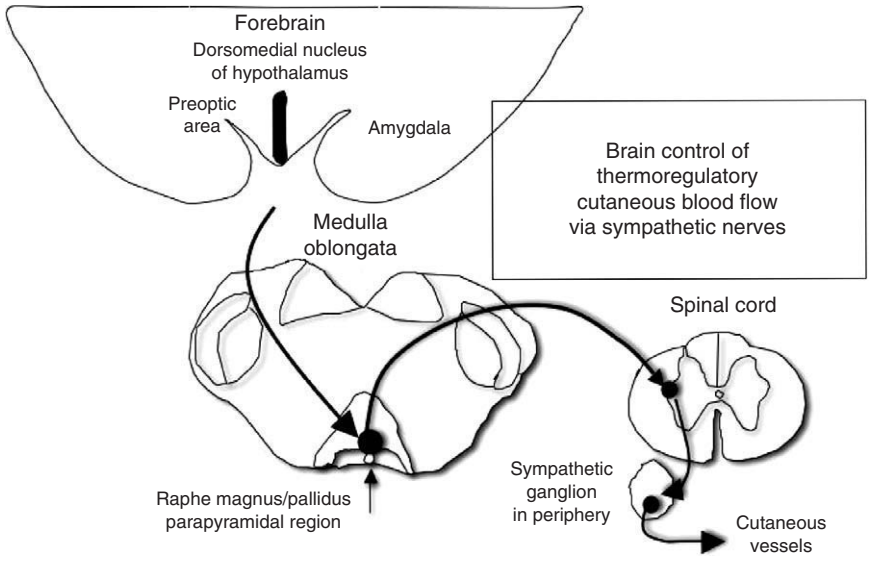


Figure 4 Summary diagram illustrating neural pathways linking brain and sympathetic preganglionic neurons in the spinal cord. The raphe/parapyramidal region of the rostral medulla oblongata is an important relay station in the descending brain pathway mediating stress-induced cutaneous vasoconstriction.

neurons present in the raphe/parapyramidal region of the medulla oblongata. Some drugs used to treat mental illness also interact with brain neurotransmitter systems (serotonin and dopamine) regulating the thermoregulatory cutaneous blood flow to reduce stress-induced sympathetic activation, thereby increasing resting cutaneous blood flow and preventing stress-induced falls in blood flow. Measurement of cutaneous blood flow provides a new biological index of stress.

Acknowledgments

Supported by the National Health and Medical Research Council (NH & MRC) of Australia.

Further Reading

Blessing, W. W. (1997). *The lower brainstem and bodily homeostasis*. New York: Oxford University Press.

Blessing, W. W. (2005). Clozapine increases cutaneous blood flow and reduces sympathetic cutaneous vasomotor

alerting responses (SCVARs) in rats: comparison with effects of haloperidol. *Psychopharmacology* **181**, 518–521.
Cannon, W. B. (1929). *Bodily changes in pain, hunger, fear and rage*. New York: Harper and Row.
Mancia, G., Baccelli, G. and Zanchetti, A. (1972). Hemodynamic responses to different emotional stimuli in the cat: patterns and mechanisms. *American Journal of Physiology* **223**, 925–933.
Vianna, D. M. and Carrive, P. (2005). Changes in cutaneous and body temperature during and after conditioned fear to context in the rat. *European Journal of Neuroscience* **21**, 2505–2512.
Yu, Y. H. and Blessing, W. W. (1997). Cutaneous vasoconstriction in conscious rabbits during alerting responses detected by hippocampal theta-rhythm. *American Journal of Physiology* **272**, R208–R216.

Relaxation Techniques

W G Whitehouse, E C Orne and M T Orne
University of Pennsylvania School of Medicine, Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M T Orne and W G Whitehouse, volume 3, pp 341–347, © 2000, Elsevier Inc.

Progressive Relaxation
Self-Hypnosis
Autogenic Training
Meditation
Biofeedback-Assisted Relaxation
Conclusion

Glossary

Behavior therapy	A system of psychotherapy that views mental disorders as products of maladaptive learning that are best treated by techniques derived from classical and instrumental conditioning principles.
Systematic desensitization	A technique used in behavior therapy to eliminate specific fears by requiring the patient to imagine different situations involving the feared stimulus while remaining relaxed. To avoid overwhelming the patient, the imagined situations are often mastered successively in an anxiety hierarchy, from least to most fear-provoking.

Progressive Relaxation

For centuries, the prescription rest and relaxation has been a popular treatment for a wide range of maladies. The apparent curative properties of rest were being endorsed by physicians and healers long before a scientific basis for the association between stress and disease had been established. A seminal contribution to understanding the pathophysiology of stress was the 1929 publication of Edmund Jacobson's *Progressive Relaxation*, an empirically based treatise on the role of neuromuscular tension in various somatic, neurological, and psychiatric disorders. In this work, Jacobson noted that many instances of physical and mental disease include among their symptoms the failure to relax – an ability that is often regained following recovery. Using electrophysiological measures, Jacobson confirmed the presence of neuromuscular tension in various disorders and was able to document the power of relaxation to dissipate stress-related symptoms. He also reported that mental operations, such as imagination, emotion, and recollection, were expressed in muscle contractions, a finding that proved to be an important milestone in the development of the field of psychosomatic medicine. Thus, for the first time, there was compelling scientific evidence to support the widespread belief among medical practitioners that stress and anxiety contributed to disease states and that relaxation-based therapies provided effective treatment options for many patients.

The procedures involved in progressive relaxation (PR) developed directly from Jacobson's scientific studies of muscular tension. A fundamental tenet of this approach is that tension and deep muscle relaxation are physiological opposites that cannot coexist. Thus, to the extent that neuromuscular tension underlies a particular disorder, relaxation should be an effective remedy. Moreover, PR allows patients to become actively involved in their own treatment, a feature that Jacobson regarded as an important psychological benefit of the procedure. The broader objective, however, was to teach patients to incorporate relaxation skills into their lifestyles so that they could become better able to respond appropriately

to future stressors by controlling their tendency to become tense.

Technique

A preliminary condition of training is the use of responses that allow the patient to recognize muscle contraction (i.e., tension) wherever it is present. The procedure requires the patient to alternately tense (i.e., contract) and relax (i.e., extend) each of the major muscle groups of the body. The goal of this exercise is to help the patient to discriminate degrees of muscular tension, which can be controlled as warranted or, in extreme cases, supplanted with deep relaxation.

PR attempts to cultivate whole-body relaxation. Patients are encouraged to assume a comfortable position and relax. Then they are instructed to clench the left or right fist as tightly as possible and to notice the tension as it creeps up from the fist to the hand to the forearm. This is followed by instructions to relax the fist and to pay particular attention to the difference they feel between the tensed and relaxed states. The therapist's instructions merely guide the patient's attention to actual experiences that accompany variations in muscle tension; there is no reliance on suggestive techniques, such as hypnosis. The procedure is repeated using the opposite fist, then the upper arms, the facial muscles, the chest, the stomach, and the lower back; it culminates in the muscles of the lower extremities. Each time the patient is encouraged to become aware of the difference in sensation between tension and relaxation and to allow relaxation to come to predominate.

As originally outlined by Jacobson, the procedure involved upward of 100 training sessions, many focusing on single muscle groups. Pragmatic concerns soon prevailed and abbreviated versions were developed, beginning with the efforts of psychiatrist Joseph Wolpe, who adopted relaxation as a means to countercondition anxiety responses elicited in the context of his behavior therapeutic method of systematic desensitization. Contemporary approaches to PR attempt to promote profound whole-body relaxation in a single session. They routinely involve the alternation between tensing and relaxing groups of muscles, but the procedures take place in a matter of minutes rather than days and they frequently include hypnotic-like suggestions to facilitate relaxation.

Clinical Effectiveness

PR has been applied clinically to various medical disorders in which indicators of stress, such as sustained autonomic arousal or chronic muscle tension, are present. Both controlled and uncontrolled studies support its effectiveness in the treatment of anxiety disorders, depressive symptoms, anticipatory nausea associated with chemotherapy, tinnitus, insomnia, low back pain, hypertension, and tension headache. The evaluation of the specific contribution of PR in treating these conditions is often complicated by the inclusion of other potentially therapeutic components in the treatment protocols. At the same time, the case for some disorders (e.g., hypertension) generally is inconclusive, owing to the fact that, on an individual basis, stress may not be a primary etiological factor. As Herbert Benson and his colleagues have noted, relaxation can lower an individual's blood pressure, but the magnitude of the effect depends on the extent to which stress plays a role in the first place. A number of studies reported in the literature on the clinical effectiveness of relaxation techniques have failed to make this important initial determination.

Finally, classical PR may not be appropriate for certain disorders, due to its potential to exacerbate symptoms. This has been reported, for example, among individuals suffering from tension headache or myofascial pain disorders, as well as with the elderly, in whom the practice of alternately tensing and relaxing relevant muscle groups actually increased pain and anxiety. Fortunately, when such complications arise, PR can be modified to omit the muscle contraction component of the exercises, which Jacobson recommended only as a means to sensitize patients to the presence of tension in their bodies. In some studies, imaginal relaxation, in which participants are not required to enact muscle tension–release exercises but merely to imagine performing them, was associated with reductions in stress symptomatology. In addition, recent efforts to customize relaxation therapy to the functional abilities of individuals with spinal cord injuries, as determined by the level or site of injury, have produced positive outcomes in several studies.

Self-Hypnosis

The use of hypnosis in stress-related conditions is another technique that seeks to engender deep relaxation, typically with the use of direct suggestions that the subject experience calm, sleepiness, and contentment (see **Hypnosis**). A preponderance of data shows that virtually all hypnotic phenomena that can be experienced in a dyadic relationship with a therapist can also be induced by the patient using self-hypnosis. This is because heterohypnosis is fundamentally self-hypnosis, and vice versa. That is, the experience of hypnosis requires an appropriately hypnotizable individual who is motivated to accomplish the goals targeted by select therapeutic suggestions administered

in a distinctively hypnotic context. By the same token, self-hypnosis can be conceptualized as an extension of heterohypnosis in which the patient carries out a procedure taught in a therapeutic relationship and is able to reinstate feelings that were originally established in heterohypnosis by imagining the words of the therapist during the self-hypnotic exercise.

Self-hypnosis is best learned with the help of an experienced hypnotherapist who outlines the clinical objectives, dispels any misunderstanding about the procedure and its outcomes, teaches techniques to induce the condition of hypnosis, and works with the patient to establish appropriate suggestions and metaphors to facilitate therapy. Routinely, skill in self-hypnosis develops as the patient practices the techniques outside the therapist's office. To encourage regular practice and to optimize the therapeutic value of self-hypnosis, however, it is helpful for the patient to maintain his or her relationship with the therapist. We have found, for example, that in the context of stress and pain management, periodic therapy sessions, or even occasional telephone contact with the therapist, can sustain the patient's motivation to practice self-hypnosis and may even compensate for deficiencies in hypnotic ability that might otherwise produce discouraging outcomes. Attention to motivational issues is, therefore, an important aspect of preparing patients for positive experiences with self-hypnosis.

In addition to the capacity of self-hypnosis to promote deep relaxation, many individuals who practice the technique as a complement to an ongoing therapeutic relationship are also able to favorably modify their perceptions regarding the importance of the stressor, as well as their emotional reactions to it. Accordingly, for patients with sufficient hypnotic ability, self-hypnosis might well be the treatment of choice for conditions (e.g., chronic pain) that fail to respond adequately to relaxation-oriented therapies alone.

Autogenic Training

Closely related to self-hypnosis is the formalized system of auto-suggestive therapy known as autogenic training (AT). It was developed by Johannes Schultz as a form of psychophysiological therapy in which patients can be taught a series of exercises to induce in themselves a profound state of physical and mental relaxation similar to deep hypnosis. The method grew out of Schultz's own clinical work with hypnosis, during which he noted that hypnotized patients commonly reported the sensation of heaviness in their limbs, accompanied by feelings of warmth. These observations were taken and interlaced with his belief that hypnotic suggestions were successful to the extent that the patient allowed them to happen. Specifically, in Schultz's analysis, hypnotized patients adopted a type of passive concentration, permitting the suggested changes to occur without trying to directly influence their occurrence. Hence, the blueprint for AT evolved. The purpose was (1) to achieve the psychophysiological effects typical of hypnosis without the need to formally induce the condition of hypnosis and (2) to train the patient to self-administer the relevant suggestions, thereby conveying a substantial element of responsibility for treatment from the therapist to the patient.

Technique

AT can be carried out on an individual or group basis, although individual training can more readily adjust to idiosyncratic differences in rate of skill acquisition or other needs. Environmental considerations include a slightly darkened room free of distractions, with a comfortable temperature and a chair, couch, or mattress. The verbal formulas developed by Schultz and used clinically for nearly three-quarters of a century fall into three categories of autogenic exercise: standard, meditative, and special. For illustrative purposes, we outline here only the standard autogenic exercises, whose mastery is a prerequisite for advancement to meditative (involving imagery) and special (designed to treat specific organic or mental disorders) exercises. Moreover, the standard exercises are the most widely used for stress management.

- Heaviness. Training in heaviness usually begins with the dominant arm, with the expectation that, as the exercise proceeds, heaviness will begin to generalize to other extremities. The trainee passively concentrates as the therapist repeats calmly, several times, "I am at peace.... My right arm (if that is the dominant arm, otherwise, left arm) is heavy.... My right arm is heavy.... My right arm is heavy." As the right arm becomes heavier, the trainee begins to notice a sense of heaviness in the left arm as well. This is reinforced as the therapist slowly repeats, "My left arm is heavy," a number of times, followed by "Both of my arms are heavy." The exercise is then extended to induce the spreading heaviness in each of the legs, concluding with the formula "My arms and legs are heavy."
- Warmth. The second standard exercise is intended to induce the experience of warmth in the arms and legs by increasing blood flow in the extremities. Once heaviness has been achieved, the second formula is combined with the first: "I am at peace.... My arms and legs are heavy.... I am at peace.... My arms and legs are heavy.... My right arm is

warm.... My right arm is warm.... My right arm is warm." The procedure continues with repetition of the appropriate verbal formulas and passive concentration as warmth develops in the other arm and in each of the legs, culminating with repetition of the phrase "My arms and legs are warm."

- Cardiac regulation. The third standard exercise focuses on the trainees' awareness of their heartbeat. Many trainees do not readily perceive their heartbeats unless they are hyperaroused. Such individuals may require postural alterations to sensitize cardiac awareness, such as practicing autogenic exercises with the right hand placed over their heart. The relevant verbal formula that is the focus of passive concentration is "Heartbeat calm and regular." The objective is not to slow the heart rate but to bring it into a consistent rhythm that supports the already established feelings of heaviness and warmth.
- Respiration. The fourth exercise is intended to teach the self-regulation of respiratory functions. Like the cardiac regulatory exercise that precedes it, the trainees' respiration rates should have been substantially decreased during the heaviness and warmth exercises. However, unlike heartbeat activity, the rate of breathing is relatively easy to modify voluntarily, an inclination that must be resisted if passive concentration is to be maintained. Accordingly, the fourth formula is phrased "It breathes me" to emphasize the importance of passive focus.
- Abdominal warmth. The fifth standard exercise attempts to induce a generalized feeling of warmth deep in the region of the solar plexus. Many trainees require an anatomy lesson to localize this target of passive concentration (i.e., halfway between the lower segment of the sternum and the navel) before the therapeutic phrase "My solar plexus is warm" can be employed effectively. The exercise is contraindicated for people with medical conditions involving the viscera of the peritoneal cavity (e.g., diabetes, gastric tumors, and ulcers).
- Coolness of the forehead. The final standard autogenic exercise stems from the common observation that a cool cloth applied to the forehead reduces tension – an effect that can also be achieved by a redistribution of the blood flow away from the head (i.e., localized vasoconstriction) and toward the extremities (i.e., peripheral vasodilatation as accomplished in exercise 2). The relevant supplemental autogenic formula to be repeated at this stage of training is simply "My forehead is cool."

Training is considered complete when the trainee is skilled at the self-administration of the autogenic program. Each of the standard exercises works toward reinforcing the others to inhibit indicators of sympathetic nervous system activation, thereby promoting a generalized state of physiological and mental relaxation.

Clinical Effectiveness

AT is widely used in conjunction with biofeedback techniques to promote stress reduction, although many clinicians employ autogenic therapy as a solitary method to achieve self-regulation of the autonomic nervous system. Attempts to assess the specific effectiveness of AT by reviewing the relevant scientific and clinical literatures are, however, impeded by several factors (e.g., small nonrandom samples, poor controls, and wide variation in training protocols and extent of training). Quantitative reviews that ignore the idiosyncratic aspects of individual studies and focus, instead, on the magnitude of the treatment effects across studies, compared to no treatment or to other nonpharmacological techniques (e.g., biofeedback, meditation, progressive relaxation, or self-hypnosis) tend to conclude that AT is as efficacious, on the whole, as other biobehavioral interventions for most psychosomatic disorders. There is, however, some evidence that AT may specifically promote decreases in heart rate that are independent of the general lowered respiratory frequencies associated with other relaxation procedures such as Jacobson's PR technique. Stress-related conditions that are claimed to respond positively to AT include respiratory problems (e.g., bronchial asthma) and circulatory disorders (hypertension, Raynaud's disease, tachycardia, and cardiac arrhythmia), insomnia, and anxiety.

Meditation

Meditation refers to a collection of practices, generally of Asian origin, that induce a change from the ordinary in a person's mental focus, with the result that metabolic activity is slowed and the practitioner often feels relaxed and refreshed. Two general forms have been identified, concentrative and nonconcentrative methods.

In the concentrative approach, typified by Transcendental Meditation, the individual is seated in a quiet environment and directs his or her focus on a single repetitive stimulus, such as a word or sound recited mentally as the person takes in or exhales breaths of air. The objective is to remain single-minded. If distracting thoughts intrude, the meditator is directed to simply dismiss them and to bring his or her attention back to the focal stimulus. The technique, along with its religious and philosophical trappings,

was popularized in the western hemisphere in the late 1960s by the Maharishi Mahesh Yogi, whose disciples included many celebrities of the period.

A prominent form of nonconcentrative meditation in use medically is mindfulness meditation, which is derived from Buddhist tradition. The technique does not necessarily lead to a profoundly relaxed state, as is typical of concentrative forms of meditation. Rather, the primary goal is to obtain insight into the self by learning to catalog moment-to-moment changes in experience.

Mindfulness meditation begins with a single object of focus (e.g., breathing) to establish calmness, but the mental focus is gradually expanded to include any ambient stimuli, thoughts, feelings, and physical sensations that enter awareness. Unlike concentrative meditation, practitioners of mindfulness do not regard these experiences as distractions to be ignored but instead dispassionately focus on each until their attention wanders, whereupon they simply note where their thoughts have taken them before redirecting attention back to their in-the-moment experiences. Various formal and informal exercises have been developed to facilitate the adoption of mindfulness as a lifestyle approach to foster an appreciation for the importance of living in the present moment.

Clinical Effectiveness

A number of scientific investigations of Transcendental Meditation were undertaken in the early 1970s by Robert Wallace and Herbert Benson, which documented the ability of meditation to induce deep relaxation and to oppose sympathetic arousal. Among the physiological changes produced during meditation were substantial decreases in the rate and volume of respiration, with correspondingly reduced levels of oxygen consumption and carbon dioxide production. In addition, a slowing of the heart rate, accompanied by increases in skin resistance and decreases in blood lactate levels signified an overall reduction in anxiety or arousal. Furthermore, studies using electroencephalograms (EEGs) confirmed the presence of a low arousal state – a pattern featuring a predominance of alpha activity, with occasional brief transitions toward theta and deep-sleep-like delta frequencies. What was not found in these early studies was evidence that meditation significantly affected blood pressure, which Benson and his colleagues soon realized was due to the fact that their sample of experienced meditators had quite low blood pressure to begin with. Subsequent research determined that meditation was effective in reducing blood pressure in individuals with stress-related hypertension. Recent studies have also found that meditation is associated with improvements in respiratory and cardiovascular functioning, sleep disturbance, and depression, as well as in the treatment of addictive disorders.

As mentioned previously, nonconcentrative forms of meditation, such as mindfulness, do not uniformly induce relaxation. In fact, although relaxation states often occur when practicing the technique, because the intended goal is greater self-awareness no particular physiological outcome should be anticipated. Nevertheless, mindfulness meditation appears to be notably useful for the management of chronic pain, in which learning to identify the everyday pain sensations apart from their emotional and evaluative components helps to lessen the degree of suffering that is otherwise endured. Preliminary evidence suggests that training in mindfulness meditation may also be of benefit in women with breast cancer by reducing stress and anxiety and improving their outlook for recovery.

Biofeedback-Assisted Relaxation

Many of the bodily processes affected by relaxation techniques are normally unavailable to conscious awareness; hence, they elude any casual attempts at voluntary control. Biofeedback rectifies this situation through the use of sensors and instrumentation capable of detecting, amplifying, and displaying the relevant biological signals. Armed with information regarding the moment-to-moment status of their blood pressure, heart rate, muscle tension, or electrical activity of the brain, patients are in a position to observe any correlations between thoughts and behavior and corresponding changes in these physiological modalities. Once such relationships are discovered, it becomes possible to achieve self-regulation of these systems with the help of biofeedback. In many cases, this self-regulation can be maintained outside the therapist's office (i.e., without reliance on instrumentation) by carrying out the cognitive or motor strategies that were previously demonstrated to affect relevant aspects of the person's physiology.

Biofeedback is used in the treatment of many disorders, some with central nervous system involvement (e.g., cognitive and attentional problems, and depression), others involving the circulatory system (e.g., cardiac arrhythmia, hypertension, and orthostatic hypotension), and others involving localized neuromuscular dysfunction (e.g., urinary or fecal incontinence and tension headache). As an aid to relaxation, however, two modalities are generally preferred: electromyographic (EMG) feedback and temperature feedback.

Electromyographic Biofeedback

EMG biofeedback is employed to monitor and display muscle tension. Although almost any muscle that can be monitored with skin surface electrodes is a viable target for EMG biofeedback, three muscle groups tend to be favored: the frontalis muscle (forehead), the masseter (jaws), and the trapezius (shoulders). The reason these muscles receive special attention is that they are particularly prone to contract during stressful situations. The goal of EMG training is to use the biofeedback information to learn particular cognitive or behavioral strategies that reliably induce relaxation in the targeted muscle site. EMG biofeedback is designed to accomplish the same objectives as Jacobson's progressive relaxation technique while permitting the patient to observe his or her progress.

Temperature Biofeedback

Thermal biofeedback is used to relay information about skin temperature to the patient. To accomplish relaxation, the patient is trained to increase the skin temperature of the hands and/or feet. The instrumentation can be quite simple (e.g., an outdoor thermometer taped to a finger or toe is usually adequate), making it convenient for patients to practice the technique at home. The physiological mechanism that supports the intended alterations in skin temperature is an increase in blood flow in the extremities, which opposes the usual sympathetic response to stress, involving peripheral vasoconstriction.

Clinical Effectiveness

Biofeedback-assisted relaxation is the most commonly employed application of biofeedback techniques. EMG biofeedback procedures have been found to be effective in the treatment of anxiety, insomnia, and various psychosomatic disorders, such as colitis, hypertension, gastric ulcers, and tension headaches. Thermal training has been particularly helpful in the treatment of migraine headache and Raynaud's disease (a vascular disorder involving the sensation of extreme cold in the hands and feet). It is frequently combined with autogenic phrases, imagery, and breathing exercises.

Conclusion

Whether provoked by palpable adverse life events, microscopic pathogenic agents, or simply our own misperceptions and anxious thoughts, stress is an inescapable part of human experience that demands biological and behavioral adjustments. Unfortunately, when exposed to persistent stress, constitutionally vulnerable individuals or people with inadequate coping resources may suffer harmful health and psychological consequences. The development of relaxation and stress-reduction skills provides an often pleasurable and cost-effective approach to moderating the experience of stress before it takes a serious toll. An important caveat, however, is that each of these techniques has the potential to modify physiological and psychological processes in beneficial and sometimes harmful ways. It is, therefore, strongly advised that individuals motivated to develop relaxation skills consult with a qualified physician, psychologist, or dentist who can recommend an appropriate program that is, at the same time, suited to the person's lifestyle. The most significant benefits of relaxation techniques usually accrue when practice becomes part of a daily routine.

See Also the Following Articles

Anxiety; Cytokines, Stress, and Depression; Depression and Manic-Depressive Illness; Depression Models; Hypnosis.

Further Reading

Benson, H. and Klipper, M. Z. (2000). *The relaxation response*. New York: Harpertorch.

Davis, M., Eshelman, E. R. and McKay, M. (2000). *The relaxation & stress reduction workbook* (5th edn.). Oakland, CA: New Harbinger.

Kabat-Zinn, J. (1991). *Full catastrophe living: using the wisdom of your body and mind to face stress, pain and illness*. New York: Delacorte.

Lehrer, P. M. and Woolfolk, R. L. (1993). *Principles and practice of stress management* (2nd edn.). New York: Guilford.

Religion and Stress

S Packer
New School for Social Research, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Packer, volume 3, pp 348–355, © 2000, Elsevier Inc.

Religion as a Concept
Religion as a Cure for Stress
Religion as a Cause of Stress
Religion as a Correlate of Stress

Glossary

Apocalyptic religion	A religion based on writings prophesying a cataclysmic time when evil forces are destroyed; includes millenarian religions that focus on the end of 1000-year period described in the Book of Revelation.
Cult	A pejorative term used to describe a new religious movement, especially one that is secretive, is socially isolated, is faddish, has unfamiliar rules and rituals, and uses coercive or manipulative or deceptive techniques used to win or keep new converts.
Prophetic religion	A religion that maintains that truth is revealed by prophecy; sometimes referred to as religions of revelation. Prophetic religions include Western monotheistic religions such as Judaism, Christianity, and Islam, as well as Zoroastrianism and Bahai.
Religion	A system of beliefs centered around the concept of a supernatural being or force.
Transcendent religion	A religion that teaches techniques for transcending or transforming the reality of the senses or that denies the existence or importance of mundane reality. Transcendent religions include many Eastern religions and some Western New Age beliefs.

Religion can be both a cause of stress and a cure for stress. Some go so far as to say that it is religion's ability to relieve the very stress that it induces that is partly responsible for its tenacious hold on its truest or newest believers. It has also been noted that religious sentiment and sectarianism rises during times of increased personal or societal stress. Because change and uncertainty are potent causes of stress, on a psychological, sociological, and physiological level, it is not surprising that some of history's most unusual religious movements have sprung up during the times of rapid social change and uncertainty. Individuals undergoing personal stress and lifestyle shifts are statistically more likely to get involved with unusual or innovative religious movements. Thus, adolescents who have recently left home, the newly divorced or bereaved or relocated, and prison inmates are especially susceptible to the appeal of new religions or radical religions. But charismatic religious groups or cults are not the only religious groups that gain momentum during times of stress; even the ranks of more garden-variety religions grow when the stresses of poverty, old age, war, illness, isolation, or impending death increase.

Religion as a Concept

Religion is anything but a single, all-embracing phenomenon. There are over 100,000 registered religions in the United States alone. The number of registered religions is increasing, not decreasing, and has been on the rise in the United States since the 1960s. The number of people who accept a personal spirituality, rather than a recognized religion, is rising even faster, to the point that spirituality is now a more politically correct and more all-embracing term that bypasses the conservative connotations of the word religion.

Even the meaning of the word religion is a source of contention. For the sake of simplicity, we define religion here as a system of beliefs centered around a supernatural being, power, or force. But, for the sake of accuracy, we must add that many well-respected definitions of religion that include systems of belief do not include a belief in the supernatural and that some well-known religions, such as Confucianism, are not concerned with a supernatural entity.

Given the large number of recognized religions in existence, it stands to reason that religions can differ dramatically from one another. Furthermore, some organized religions include far-flung sects and subsects that bear little superficial resemblance to the parent religion, and some (but certainly not all) religions host a wide variety of views on different issues. It may be tempting to think that religious beliefs, rites, customs, and concerns can be plotted out on a spreadsheet, so that different religions can be directly compared to one another. However, this straightforward approach of scientists leads to an inaccurate appreciation of the vagaries of religion and overlooks the fact that one religion stresses rites of worship,

another concerns itself with right thinking, a third revolves around prayer or penitence, a fourth focuses on gender differences and reproduction issues, a fifth emphasizes the afterlife, and so forth. It is unlikely that each religion has something to say about every aspect of life (or death), even though each religion represents itself as the ultimate authority on human life.

Religion as a Cure for Stress

Both its fiercest critics and its most fervent advocates agree that religion (or at least some religions) can dampen psychological stress. Freud himself, that arch-enemy of religion, admitted as much in his seminal book on *The Future of an Illusion*. Personal testimonials and clergymen's sermons about religion's potential to relieve stress are too numerous to cite here. Religion is so commonly associated with stress that it is virtually a reflex to invoke it at moments of stress. Many images of religion and extreme stress spring to mind immediately, from foot-stomping spirituals of the segregation-era South to the footsteps of the death-row dweller listening to the last words of the seventy-first psalm to the prayers mumbled by chaplains for charging troops who face death on the front. We can even consider the fearlessness of the 9/11 suicide bombers who embraced their own death as (perverse) testimony to the power of faith.

Although not everyone finds relief in religion, there are some compelling sociological statistics supporting religion's overall ability to combat stress. It is well-established that people with strong personal religious beliefs and affiliations are less likely to commit suicide, which can be seen as the ultimate and most extreme response to stress. The fact that some Islamic extremists encouraged suicide missions, in spite of Islam's strong stance against suicide, is an aberration.

Philosophical Aspects

By positing the existence of a supernatural being who actively intervenes in the natural world and who has a plan for each individual human, religion provides an automatic antidote to the stress of being alone and awash in an unresponsive and uncaring universe. Philosophers and theologians alike note that existential angst and dread have plagued twentieth-century society, at least in the West. Ever since Nietzsche declared that "God is dead" as the nineteenth century ended, and as his own symptoms of neurosyphilis progressed, unwavering faith in the supernatural melted away and slowly but surely became the foremost source of stress. Only the lurking specter of a sudden nuclear holocaust overshadowed this stress. Pastoral pop psychologists tried to reverse the tide that Nietzsche set into motion. Best-selling books, such as Paul Tillich's *Man Is Not Alone*, played on Nietzsche's angst-inducing assertion in its title and appealed to this angst. Similarly, film titles and themes recapitulate the reflexive tendency to call on religion in times of crisis. For instance, the film *God is My Co-Pilot* appeared during World War II. More veiled religious themes bubble to the surface during other times of stress. When the Great Depression cast its shadow, *Lost Horizon* evoked legends of Shangri-La, where Tibetan monks and Christian missionaries fused good works and good karma and reinfused human souls. Then, as the twentieth century came to a close, and as second millennium neared – at the very same time that information technology zoomed ahead and rocked standard channels of communication – religious imagery once again resurfaced in secular guise, this time in the form of angels. New Age angels, which were retrieved from Renaissance-era paintings, appeared on postage stamps.

Psychological Aspects

Antidote to aloneness, abandonment, and dependency
When psychoanalysts try to explain why a belief in the supernatural relieves stress, they look for more personal, rather than societal, explanations. The consistently controversial Freud claimed that a belief in God substitutes for the presence of an all-protective, albeit sometimes punitive, parent. According to such theories, religion recreates the reassuring world of dependency that an infant enjoys. The oceanic feeling of connectedness with the universe that mystics describe could be a way to return to the peacefulness and protectiveness of the womb, where an umbilical cord delivered never-ending nourishment and where amniotic fluid buffered the baby from unkind blows from the world around. Religious redemption concepts that revolve around the return to Eden can be compared to psychoanalytic fantasies about the return to the womb.

More contemporary neo-Freudian psychoanalytic thinkers, such as Karl Winnicot, emphasized the importance of separation-individuation as the critical step in psychological growth. Unlike more orthodox Freudians, who viewed religious reassurance as a sign of psychological immaturity and as an obstacle to self-actualization, theorists such as Winnicot felt that religious belief functions as a useful transitional object, much like a baby blanket or a teddy bear. Both remind the child of the parent's perpetual presence and assuage a child's subjective stress of abandonment during the parents' absence. There are many, many more psychoanalytic theories about the ways that religion relieves stress, each unique and interesting in

its own right and each as untestable by scientific techniques as the tenets of religion itself.

A more pragmatic and less theoretical take on religion as a surrogate dependency is found in the principles behind the Twelve-Step and Recovery Programs and Alcoholics Anonymous. These programs provide a strong social support system and also call on a Higher Power to relieve the psychological and physiological stress of alcohol withdrawal and as an aid in dealing with stresses that were previously ameliorated by addictions. Rather than admonishing their members for depending on this spiritual Higher Power, as more Freudian-influenced therapists did in the past, the Twelve-Step groups see spiritual dependency as far less harmful than chemical dependency. Such pro-spiritual approaches have gained increasing public and professional acceptance and have been adapted to the treatment of many other disorders.

Decreased sense of randomness and uncertainty There are many other ways that religion relieves personal stress. Religion decreases the sense of randomness, uncertainty, and chaos in the world because it details the logic (or illogic) behind a world plan. It predicts or sometimes prophesies or, at the very least, attempts to explain those events that seem most inexplicable. In doing so, it creates a concept of stability that has been described as an ordered universe. Even religions that foretell the coming of adverse events in the future, such as an impending apocalypse, reincarnation, predestination, or bad karma, have the potential to offer partial relief from stress, simply because they help their members prepare themselves psychologically for adverse events. Providing explanations for events, even without providing a means for altering those events, is referred to as heuristics, and this is one of the most powerful psychological tools that religion possesses. Some critics of psychoanalysis say that its appeal also rests on its heuristic value and argue that unprovable psychoanalytic explanations provide reasons for behavior rather than remedies, in the same way that religion provides reasons for cosmic events and human nature.

Religions further relieve the stress of uncertainty by providing blueprints for behavior that can theoretically change the future. By prescribing prayer, penance, codes of charity, dietary laws, sacrificial rituals, or right thought, those religions reinstill a sense of personal control. More optimistic prophesies, and promises of paradise, confer hope and escape.

Cognitive coping techniques Above and beyond any theoretical parallels with psychoanalysis, many religions also provide specific cognitive techniques that are useful in coping with stress. For instance, the mere act of acknowledging and articulating the experience of stress can quell some distress. Long before Freud discovered his talking cure, religious prayers and petitions provided collective voices for stress and distress. Hebrew psalms that begin with words such as "From the depths I called out unto Thee" poeticize an acute sense of anguish, whereas Christian recreations of the crucifixion scene dramatize the ordeal of Christ on the cross. These words and images are especially appealing to people who do not have access to other avenues or expression, or who prefer to deflect their own subjective sense of distress by focusing on more universal, cosmic, or collective stresses.

Role models for stress endurance Religious lore offers as role models people who withstood extreme stress and either surmounted that stress, were valued because of their ability to endure stress, or went on to fulfill higher purposes because of those stresses. The Christ-figure is the consummate example of this sort, but it is hardly the only one. Biblical stories about Jonah being swallowed by a whale, Noah's ability to withstand a flood, Daniel's survival in the fiery lion's den, and Job's endurance of the loss of his family and his health all reassure believers that there is a relief for stress or, if there is no immediate relief, that there is a reason or perhaps even a reward in another life. The martyrdom of Christian saints before their beatification, the Buddhist *jataka* stories about Prince Shakyamuni's wandering as a mendicant monk before he achieved enlightenment as he sat under the Boddhi tree, and the image of the Israelite tribes wandering through the desert for 40 years before entering the Land of Israel and acquiring their status as the chosen people are other examples of the pivotal role that stress endurance plays in religious themes.

Consolation, devaluation, and dissociation Some religions provide such a sense of consolation that they have been dubbed the religions of consolation. Christianity's contention that the meek with inherit the Earth reassures people that their this-worldly stress and suffering will be relieved by other-worldly rewards. The promises of future paradise made by prophetic religions dull the pain of the present and appeal to the impoverished, the downtrodden, and the psychologically stressed. Some religions offer so much consolation for contemporary stress and distress that they devalue the material world completely by undermining its ultimate importance or by teaching that the real world is but a delusion and no different from a dream. Buddhism's belief that the material world is nothing more than a delusion, or *maya*, is an extreme example of such transcendent thinking. The Hindu yogin's aspiration to a waking

state of dreamless sleep is another example of transcendence. Some psychotherapeutic techniques train patients to use similar techniques of detachment when confronted with stress-producing stimuli, although it cannot be overlooked that people who enter such dissociative states spontaneously and nonvolitionally develop serious difficulties in life.

Temporal and physical escapes from stress Even world-affirming religions, which affirm the importance of the material world, often offer temporary escape routes from daily stress. Religious holidays and religious services and other sacred times carve out stress-free time during the ordinary workweek and create an opportunity for rejuvenation. For those times when real-world stresses require even more relief than routine religious beliefs or rituals can confer, some religions offer physical as well as temporal retreats where food, clothing, and shelter are available, along with social support, structure, and spiritual exercises. Ashrams, yeshivas, monasteries, convents, and any number of other religious communities provide parallels to the retreats popular among certain Christian denominations. Such religious retreats not only legitimize the need for relief from stress, but also consecrate the choice to retreat from stress. Such religious retreats thereby provide participants with a renewed sense of self-worth, along with a sense of connectedness with others who share their belief system, and may even offer retraining in new vocations or avocations. Some retreats encourage members to contribute to society by doing good works or charity. The modern hospital movement evolved out of the monastic retreats in the Middle Ages, where people who originally sought relief for spiritual, physical, psychological ailments eventually provided care for others after their own recovery. In contrast, psychiatric rest cures and funny farms, which also offered retreats from the real world, stigmatize the participant, pathologize the process, and end social productivity. It is no wonder that religious retreats are often preferred over psychiatric treatment and that religion is often the first defense against stress.

Social Aspects

For some people, and for some religions, religion is a solitary matter and nothing but. For them, it is personal spirituality that matters most, with respect to stress relief and everything else. Although the title of William James's often-republished book is *The Varieties of Religious Experience,* in it James focused exclusively on religion that is experienced in solitude. Nevertheless, religion exists on a social as well as a personal level and acts to relieve social stress (or produce stress) on both levels as well. By providing a social support network through their communal services and activities, coupled with a sense of collective identity and purpose, religious organizations can directly combat the stress of loneliness, displacement, and *anomie* that Emil Durkheim implicated in his early sociological studies of suicide.

Some religions provide material benefits, such as charity, lodging, employment, social services, and subsidized medical care, which can counteract the intense stresses of economic hardship and ill health and can complement the psychological consolation inherent in religious theory. Religious bureaucracies also offer alternative, and much appreciated, channels for personal and political expression, particularly when access to legitimate political clout is closed. Moreover, organized religion can become powerful enough to challenge the existing political and economic powers and bring about lasting and legitimate social and legal change.

Studies of the conversion process are especially illustrative of the importance of social forces and religion. It often comes as a surprise to people who are ideologically committed to religion to learn that it is the social sway of a religious group that influences individuals to adopt new religious ideas rather than the other way around. In other words, rather than experiencing a life-changing epiphany before converting, similar to the epiphany that the Gospels attributed to Saul of Tarsus on the road to Damascus, converts to new religions are more likely to follow a socially paved path to new religious insights. They accept more and more of the ideas of the group as they gain greater and greater acceptance into that group and as they become more reliant on group members for social support. The process of conversion is more likely to be gradual than sudden; in most people, behavior changes incrementally rather than dramatically. Nor does everyone who gets involved with a new religious group remain committed to that group. Of the many factors that correlate with an individual's long-term involvement with a new religion, one of the most important factors revolves around the role of stress. The more stress relief that an individual experiences at the time of joining a new religious group, the more likely that person is to remain a member of that group. Furthermore, it is the stress that individuals experience when distancing themselves from such groups that typically compels them to return to those groups.

Because it is well known that people who are already in a state of stress stand to achieve the greatest degree of stress relief from a new religion, some such religions make it a point to stress their newest members. Techniques such as sleep deprivation, diet

restriction, social isolation, overwork, enforced silence, sexual abstinence, or even physical or psychological threats make new recruits more receptive to the stress-relieving effects of the religion. Religions that rely on such insidious, quasi-coercive stress-inducing techniques are often denounced as cults and are disdained by nonmembers or former members. Such cultish religions predictably recruit in institutional settings, such as colleges, prisons, and retirement communities, where people are already in high stress states. Because such residents are pre-primed for proselytization, new religions have made major inroads in such places, at times to tragic ends.

It is the intense stress of prison life, rather than an earnest desire to reform or to atone their past acts, that propels some inmates to seek relief through religious conversion. Many ministries maintain an active presence in prisons because they recognize this potential. Sometimes, religious recruitment goes hand in hand with political indoctrination and can lead to involvement in militant religious factions. Many Black Muslims and Seven Percenters, including Malcolm X, traced their conversion to their incarceration. Yet, until 9/11 made the U.S. public more aware of this potential, there was little public opposition to religious recruitment in prisons. Most people assumed that religions strive to instill socially acceptable values that will substitute for criminal behavior, aid in rehabilitation and reentry into society, and thwart future incarceration. The recruitment of prisoners to a new religion was preferred over recruitment by antisocial prison gangs. However, in our post-9/11 world, we must consider the possibility that some extremist religions can instill antisocial attitudes and push some individuals into committing extreme acts.

In contrast, there has been a great public outcry about cult recruitment on college campuses, partly because some cultish religions have the opposite effect of prison-based cult religions. Rather than aiding the recruit's reentry into society, these cults isolate students from society and abort their attempts to achieve the status of other adults in society. Some cults cut off members' contact with families and friends, and cause severe stress in their relatives as a result. Although many legitimate religions were considered to be cults in their early stages – with Methodism being the most notable U.S. example – the mere fact that a small percentage of cults have been associated with mass suicide or homicide is enough to cause public concern and to stimulate psychiatric task forces to investigate these issues. The memories of 800 deaths in Jonestown, the loss of 19 lives in California's Heaven's Gate suicides, the subway poisonings by Japan's Aum Shinrikyo sect, and the murder of a Hare Krishna defector in North Carolina tend to overshadow studies that show the psychological benefits of conversion in some individuals.

Somatic Aspects

Stress is as much a physiological response as a psychological one. Some religious rites use physical means to achieve higher spiritual states and alleviate both psychological and physiological stress in the process. The muscle stretching systems of yoga, the breathing exercises of Zen, and the controlled movements of Tai Chi are just a few examples of Eastern-influenced mind–body methods that gained popularity in the West. Some of these methods are now taught in continuing education courses sponsored by the American Psychiatric Association.

Although it is impossible to verify the existence of higher spiritual states, it is possible to measure changes in heart rate, respiration, galvanic skin response (sweating), pupil size, secretions of stress hormones such as cortisol or gastric acid, brain waves (through an electroencephalogram, EEG), and muscle tension in people performing those spiritual exercises. Internist Harold Benson's studies of cardiovascular effects of Transcendental Meditation and Tibetan Buddhist chanting appear in juried medical journals and in his popular paperback, *The Relaxation Response*. Japanese psychiatrist Tomio Hirai correlated the EEG and electromyogram (EMG) effects of both Zen meditation and Hindu yoga with reports of spiritual and psychological states and published his results in the curious book *Zen Meditation and Psychotherapy.* These works represent serious research, but there have also been many hyperbolic claims about the benefits of Eastern (and Western) religious systems, leading some professionals to dismiss any and all claims reflexively and prematurely. Recent scientific studies confirming the correlation between stress relief and specific religious rites suggest that both an open mind – and an open eye – can help us to appreciate confirmed the mind–body benefits of religion and also to avoid misleading claims and cults.

Quite the opposite of the calming meditative techniques of Eastern religions are the ecstatic dances, shaking, quaking, rocking, and rolling movement of some sects, which inspired the names Holy Rollers, Ranters, Ravers, Quakers, Shaking Quakers, and Shakers. Such intense activity presumably relieves stress through the same mechanisms that jogging and exercise help more secular devotees. The Dionysian dances of Classical Greece, described in Euripides' play about the Bacchae, were but one of many recurring manifestations of frenzied religious dancing that

serves a related function. Repetitive religious rituals in general are said to relieve stress as well.

Side Effects of Stress Relief

Religion's efficacy at relieving stress does not come without side effects. There are times when religion relieves stress so well that its practitioners are unaware of the demands of the real world and are unable to mount a defense against impending danger. Many social scientists have said that the religious devotion of African Americans shielded them from the pain and poverty of pre-Civil Rights America and delayed the adoption of more appropriate political tactics. Similarly, the Tibetan practice of sending one-third of its youth to Buddhist monasteries left the country defenseless against the Chinese invaders who destroyed temples, massacred monks and nuns, and sent their religious leaders into exile. Some secular Zionists claimed that the insulated and self-satisfied religious infrastructure of some eastern European Jewish communities obscured their awareness of the deadly fate that awaited Jews in Hitler's death camps. An analogous U.S. tragedy occurred when Native American warriors went into battle unarmed, believing that their Ghost Dance ritual would protect them from the guns and arrows of approaching armies. Marx and Engels blamed the mystical and occasionally bizarre religious belief of Russian orthodoxy and schismatics for numbing Russian reactions to the material exploitation by the capitalists and the Czars. This observation prompted Marx to coin his oft-quoted characterization of religion as "the opiate of the masses."

Religion as a Cause of Stress

As effective as religion can be at relieving stress, it can also produce stress. Threats of an afterlife full of hellfire and brimstone or of an impending apocalypse that will destroy the world are obvious stress-producers. On a more subtle level are the many moralistic demands and behavioral codes made by religion, which are often difficult to live up to and, thus, tend to leave some practitioners in a near-constant state of imperfection and incompleteness. Some practitioners adopt even more zealous beliefs and behavior in order to avoid that stress, in an ever spiraling pattern. Skeptics such as psychiatrist John Sargeant observed that some charismatic religions exploit this tandem stress relief–stress production effect when they proselytize and compared this push–pull effect of religion to behavioral conditioning, more nefarious methods of mind control and brainwashing, and even drug addiction.

For sure, religion-induced stress is not limited to perceived threats, nor is it confined to the personal psyche. In spite of its promises of eternal peace, organized religion has indeed injected some of the most realistic threats the world has ever witnessed. The Inquisition, the Crusades, and the wars of religion are but a few testimonies to reality-based stresses posed by religion in the past. Such stresses persist to the present day through suicide cults such as the California-based Heaven's Gate, Islamic terrorist attacks on the World Trade Center, the biological-weapon-wielding Aum Shinrikyo in Japan, the Hindu–Muslim–Sikh conflicts in an atomic-bomb-armed India, the religious-ethnic conflicts of the Balkans that set the stage for World War I, and right-ring religious assassinations in Jerusalem, to name just a few. Although some religions aspire to the day that "the lion will lie down with the lamb," that day has not yet arrived. Religion has been, and probably always will be, as intimately associated with the stress of war and violence as with the proverbial love and peace that it promises. Rising religious fundamentalism promises to produce more political and personal stresses for individuals and for the world at large. At the same time, the religions of reassurance will provide personal stress relief for individuals, creating a never-ending seesaw.

Religion as a Correlate of Stress

The correlation between the rapid rise of new and sometimes radical religious movements and the degree of social stress is nothing less than remarkable; it is cataloged in Norman Cohn's classic *Pursuit of the Millennium* and in his more recent *Chaos, Cosmos, and the World to Come*. A more recent book, *Radical Religion in America,* zeros in on extreme religious schisms that have arisen since 1970, during the times of high stress that followed the social turbulence of the 1960s. It is noteworthy that psychoanalysis, a once-esteemed and all-embracing psychiatric treatment that survives as a mere shadow of its former self, has also been accused of being a cult and of inducing a cultlike devotion that resembles religion more than science. Note that psychoanalysis's popularity peaked in the 1950s, during the times of high stress that followed World War II.

Further Reading

Cohn, N. (1993). *Cosmos, chaos, and the world to come.* New Haven, CT: Yale University Press.

Freud, S. (1961). The future of an illusion (1927). In: Strachey, J. (ed.) *Standard edition of the complete*

psychological works of Sigmund Freud (vol. 21), pp. 3–56. New York: W. W. Norton.
Fuller, A. (1994). *Psychology & religion* (3rd edn.). Lanham, MD: Rowman and Littlefield.
Gallant, M. (ed.) (1989). *Cults and new religious movements*. Arlington, VA: American Psychiatric Association.
Girard, R. (1993). *Violence and the sacred*. Baltimore, MD: Johns Hopkins University Press.
Hirai, T. (1989). *Zen meditation & psychotherapy*. New York: Tokyo Publications.
James, W. (1982). *The varieties of religious experience*. New York: Penguin Books.
Kakar, S. (1982). *Shamans, mystics, and doctors*. Chicago: University of Chicago Press.
Kaplan, J. (1997). *Radical religion in America*. Syracuse, NY: Syracuse University Press.
Kinsley, D. (1996). *Health, healing, and religion*. Upper Saddle River: Prentice-Hall.
Klass, M. (1995). *Ordered universes*. Boulder, CO: Westview Press.
Meissner, W. W. (1984). *Psychoanalysis and religious experience*. New Haven, CT: Yale University.
Ostow, M. (ed.) (1982). *Judaism & psychoanalysis*. London: Karnac Books.
Packer, S. (1998). Jewish mystical movements and the European ergot epidemics. *Israel Journal of Psychiatry* **35**, 227–241.
Paloutzian, R. (1996). *Invitation to the psychology of religion* (2nd edn.). Boston, MA: Allyn & Bacon.
Sargant, W. (1957). *Battle for the mind*. New York: Penguin Books.
Schumaker, J. F. (1992). *Religion and mental health*. New York: Oxford University Press.

9/11, Religion and Stress

S Packer
New School for Social Research, New York, NY USA

© 2007 Elsevier Inc. All rights reserved.

The Who, What, Where, How, and Why of 9/11
Paradoxes about Post-9/11 Stress, Suicide, and Religion
Muslims and Mental Health after 9/11
Jewish Issues Associated with 9/11 Stress
Concerns of Christians and Others after 9/11

Glossary

9/11	September 11, 2001, the day that the World Trade Center was destroyed and the Pentagon was attacked by organized conspirators led in absentia by Osama Bin Laden.
Islam	A prophetic religion that is based on the teachings of the prophet Mohammad, as recorded in the Koran. (Also called Mohammadanism.) Followers of Islam are known as Muslims and may be of any race or ethnicity and may live on any continent. Many, but not all, Muslims are also Arabs.
Posttraumatic stress disorder (PTSD)	A constellation of symptoms that occur after experiencing or witnessing a life-threatening episode. People with PTSD reexperience the original trauma through nightmares or flashbacks and feel detached or numb in daily life.
World Trade Center (WTC)	A complex of office buildings (also known as the Twin Towers after the two tallest buildings) located in lower Manhattan, New York City, that housed 40,000–60,000 workers. It was destroyed on September 11, 2001 (9/11), and the site where the WTC previously stood came to be known as Ground Zero.
Zionism	A religious and political belief that the political state of Israel is the Jewish homeland. Most (but not all) religious Jews are Zionists. Some secular Jews are Zionists, and many secular Jews are not Zionists. There are many different degrees of Zionism.

An association between stress and religion has always existed, long before September 11, 2001, came to be known as 9/11 and long before the WTC was destroyed. Before 9/11, religion was most often seen as a cure for stress, but it could also be conceived of as a cause of stress and was certainly recognized as a correlate of stress. But on 9/11, the relative importance of those associations shifted. For most people in the United States, and especially for people who lived and worked in and near New York, religion suddenly shifted to being a serious source of stress. In fact, some very preliminary studies showed that people who were most religious were more likely to experience more stress after 9/11, although the reverse is usually true. In general, religious belief and affiliation offer protection against stress, but the reasons why

religion failed to protect against 9/11-related stress will become obvious during our discussion.

The Who, What, Where, How, and Why of 9/11

What was 9/11, and what made that day so significant to Americans in general and to New Yorkers in particular? On that day, the WTC was destroyed by young Islamic suicide bombers who flew two hijacked airplanes into each of the Twin Towers. At nearly the same time, another hijacked plane hit the Pentagon in Washington, D.C. A fourth passenger plane, bound for San Francisco, was also hijacked, but was commandeered by passengers and diverted from its target destination. That plane crashed outside Pittsburgh and claimed far fewer lives than it was intended to. Each of these events was part of a well-orchestrated and partly successful conspiracy to fell the United States because these Muslim extremists led by Osama bin Laden viewed the United States as an aider and abettor of Zionists who controlled the state of Israel. It should be noted, however, that this Tuesday morning attack on the WTC was not the first. Several years earlier, a van containing a bomb exploded in the WTC parking lot, killing six people and causing commercial chaos in subway-level shops; this culminated in the conviction of a blind sheikh who led this militant Islamic sect.

Some 2000 lives were lost during the second WTC attack on 9/11. Several hundred others died in the Washington area. None of the passengers aboard the hijacked planes survived. Although citizens of many cities, indeed, many countries, were affected by the tragic events of that day, 9/11 was most closely associated with New York City after the silhouettes of the WTC's Twin Towers vanished from the city skyline.

Situated in downtown Manhattan, near New York City's financial district and the world-famous Wall Street, the WTC also impacted ordinary New Yorkers who lived nearby. Those people were not necessarily involved in the financial or insurance or communications businesses that were centered in the WTC. Some New Jersey bedroom communities, short commutes from the WTC, were also impacted after dozens of their residents who worked in the Towers died that day.

The WTC was close to the bustling streets of Chinatown, a mecca for tourists and the culturally curious and home to long-time residents, recent arrivals, and legal and illegal immigrants from mainland China, Taiwan, Hong Kong, Vietnam, and Malaysia. Buddhists, Christians, Catholics, Confucians, Taoists, and Chinese ancestor worshipers coexist on the streets of Chinatown and often intermingle within a single family unit united by its shared Chinese ancestry but not split by religious affiliation, as Western families often are. The once-lively Little Italy was also within walking distance of the WTC, as was the historically Jewish Lower East Side. In Little Italy, festivals named for regional Roman Catholic saints (such as San Gennarro) rose beyond their religious origins and attracted carnival-loving crowds of all faiths. It was to the Lower East Side, nostalgically known as Delancey Street, that poverty-stricken Jews fled the pograms at the turn of the twentieth century. They turned their pushcarts into a sprawling Sunday shopping district that gradually gave way to trendy clubs with outrageous names like Lansky Lounge. The once-grand synagogues built by struggling immigrants were often in need of restoration because younger generations abandoned their grandparents' turn-of-the-century terrain and fled to suburbia. The Yiddish theaters that once lined Second Avenue below Fourteenth Street were now reduced to engraved copper stepping-stones that were embedded in sidewalks, engraved with names of yesteryears' Yiddish-speaking stars.

The largely Hispanic neighborhood of Loisada, or Alphabet City, was also a healthy walk from WTC, even if it was not near enough for lunch. Dotted with small Catholic churches, Pentecostal centers, *espirismo* storefronts, and even some *santeria* supply shops, this section was more religiously diverse than might be expected. The few blocks that were left of what were once, respectively, Little Poland and Little Ukraine lay slightly to the west of Loisada and due north of the Twin Towers, and were home to ornate Orthodox churches and even some shops catering to neo-pagan practitioners. Little India, which should rightfully have been called Little Bangladesh because its restauranteurs actually came from Muslim country of Bangladesh, was caddy-corner from Little Ukraine. And so on and so forth.

South Street Seaport, an old-time fish market that had been transformed into a mall-like tourist attraction, was an after-work refuge, replete with busy bars, shops, and eateries. Tribeca was a newly minted, mostly loft-living community on the lower west side of Manhattan; pricey, arty Soho lay a little to the east of Tribeca and the Towers. Battery Park City, with its combination of luxury condos and subsidized housing, was also a stone's throw away and was populated by young families in need of affordable housing and by financial firms that needed crash apartments for out-of-town deal makers and overnighters. The West Village was not too far away either, with its authors, artists, and actors and its ever-active gay presence, which survived and thrived in spite of the devastation of the AIDS epidemic. The East Village, which was not quite as close

to the WTC as the West Village, was home to students, musicians, aspirants of all sorts, and also addicts, to judge by the reports of drug-related arrests. Several colleges and universities dotted the lower Manhattan landscape as well, including New York University.

Paradoxes about Post-9/11 Stress, Suicide, and Religion

There was a reason why we have delved into the details of these different downtown neighborhoods and business districts. We need to understand who lived downtown, why they were affected by 9/11, and why religion impacted some more than others. Studies have shown that the closer a person lived to Ground Zero, where the WTC once stood, the greater his or her chance of developing PTSD. That finding is not so surprising, considering how devastated the landscape was and how disturbing it was to view the carnage on a day-to-day basis and to smell smoldering flesh for weeks or maybe months later. What was surprising was that some studies showed that people with strong religious beliefs were more likely, not less likely, to experience stress post-9/11, even though religion ordinarily provides protection against stress. There were reasons why religion produced stress in this situation.

Other surprises awaited researchers after 9/11. Significantly, it turned out that people who received emergency debriefing were more likely, not less likely, to suffer long-lasting psychological symptoms of PTSD. It is now believed that rehearsing the traumatic event consolidated bad memories and made them more permanent than they might have been had the memories simply been allowed to fade away over time. An even bigger surprise concerned the prevalence of PTSD overall; post-9/11 PTSD was not nearly as widespread as first predicted.

Perhaps the biggest surprise overall concerned the role that religion played as a promoter of, rather than a protector against, suicide. Muslims, who ordinarily oppose suicide, might have even more reason for such surprise. Before we examine any specific associations between religion and suicide, let us look at some general associations between stress and religion.

Religion is often sought out in times of stress. It typically acts as a psychological salve. Repeatedly, religion has been shown to diminish the chances of completed suicide, which is the ultimate measurement of stress's effects. New religions are also known to thrive during times of social stress; recruitment to cults is especially successful during times of rapid social change. Moreover, religious participation generally decreases both criminal activity and criminal recidivism (although high-profile exceptions to this trend have admittedly come to light).

Yet the events during and after 9/11 contradicted some of these basic assumptions (and I deliberately used the word assumptions because some of these givens about religion are not necessarily historically factual). For one thing, the mere fact that the suicide bombers committed both suicide – and homicide! – in the name of religion seems to stand established wisdom on its head. After all, study after study shows that religious faith and affiliation protect people from suicide and help people through severe depression, financial devastation, even grave and disabling illness. Religion as a generic concept is supposed to instill a reverence for life, not a disregard for life (although some Eastern religions devalue daily life to the point that they dismiss everyday existence as *maya*, or illusion, and some fundamentalist religious sects place more value on the heavenly afterlife rather than on the earthly this-life).

Yet here was a situation in which the strong religious beliefs of the suicide bombers not only did not protect them from suicide but actually pushed them to suicide. Moreover, their religious belief promoted not only suicide but also homicide – and promoted not just homicide but mass homicide. On the surface, the bombers' behavior appears paradoxical. No wonder some studies showed that people with strong religious beliefs felt more stressed after 9/11 than those with lesser religious commitment. Their confidence in the protective effect of religion was challenged. Any preexisting assumptions about the automatic ability of religion to counteract impulses to hurt oneself or others could no longer be accepted as absolute fact.

If we hypothesize that at least some of the people with strong religious adherence had gravitated to religion because they needed an extra buffer against stress in the first place, then it becomes even easier to appreciate why many felt more stressed after the 9/11 suicide attacks. If their psychological foundations were already shaky, but were buttressed by religious belief that now wobbled, then they would feel shakier still. They might need to distance themselves from the perpetrator religion (Islam), to prove that their own religion still provided them protection. It would be to their psychological advantage to deride Islam at face value, to draw a distinction between their own religion and Islam, and even to wage war against Muslims in order to protect their own psychological turf and to fortify their own mental foundation.

To understand the motives of these Muslim suicide bombers, we must realize that they acted on the earnest belief that they would earn themselves and their

families eternal paradise for their valiant acts of self-sacrifice. That reward system stands in direct contradiction to the hellfire and brimstone that presumably awaits devout Christians who choose suicide. The social prestige provided by those Islamic suicides is also the polar opposite of the social ostracism faced by Jewish families whose relatives elect to suicide. Because Jews who commit suicide cannot be buried in a Jewish cemetery, their survivors are deprived of the comfort of a collective mourning ritual, and the individual is deprived of an identity and a commemoration after death.

However, it should be said that 9/11 was by no means the first time that suicide was committed in the name of religion. Nor was Islam the first religion to encourage suicide in the name of faith. Christ's willing self-sacrifice on the cross could arguably be said to be a sort of self-accepted suicide. It is less controversial to claim that the desert saints who starved to death in the third and fourth centuries were also committing suiciding in the name of faith. Many Christian martyrs have been beatified because of their willingness to endure torture unto death in the name of religion. The Hindu *sati*, a cultural *cum* religious rite that requires the widow to climb on to the funeral pyre of her deceased husband and to join him in cremation and reincarnation, was a once-respected (but now suspect) religious rite of self-sacrifice. Jews who chose suicide over surrender at mountaintop fortress of Masada or who refused to convert to Christianity when the Spaniards offered them a choice between the cross or the Inquisition could be said to have chosen suicide, albeit under coercion. This was quite different from the Muslim suicide volunteers who went to their deaths willingly and equally willingly agreed to take the lives of many others.

Suicide through self-starvation is an acceptable, even respectable, end for some ascetic sects of Jains in India. In the Middle Ages, the Cathars practiced self-starvation as a religious rite. And let us not forget the soldiers who volunteered to serve in wars based on religion knowing that they might not survive the battle but believing in the utility of their self-sacrifice.

9/11 challenged the standard associations between stress and religion in other important ways as well. As previously mentioned, recruitment to new religions or cults typically peaks at times of social stress, as does religious involvement in general. But there is no evidence that such a trend took place in the wake of 9/11 because of very specific interventions. If anything, recruitment to psychiatric treatment peaked instead, as prevention efforts to avert the predicted outbreak of PTSD diverted interest away from religion and into more mainstream therapeutic approaches. Public health and social service organizations made massive outreach efforts in the downtown area. Emergency debriefing stations were set up in lobbies in lower Manhattan, where residents were informed about both public and private psychiatric services. The Red Cross subsidized psychiatric treatment, even for those who could otherwise afford it. Posters recommending mental health care appeared in subway cars and at bus stops and reminded commuters that mental health care was not only easily available but was also readily affordable. As it turned out, these seat-of-the-pants psychiatric treatments did not accomplish what they purported and now seem to have increased the stress symptoms and PTSD. But they may have prevented religious-cult recruitment in an otherwise susceptible population.

For anyone who came of age at a time when public service announcements reminded Americans that the "family that prays together stays together" and when religion was the first referral made to someone suffering from stress or distress, seeing these subway posters was startling. These posters signaled a significant shift in social values as well as in clinical care. We might say that making a first-line recommendation for psychiatric care, rather than for religious rituals, was a sure sign that psychiatric treatment had finally come of age and was de-stigmatized enough to appear in public places outside of student counseling services.

Another explanation for the strong emphasis on mental health and PTSD in the immediate aftermath of 9/11 may have reflected the professional make-up of New York City. New York City is home to an unusually high number of mental health and social service professionals whose services are available at a moment's notice. Plus, psychotropic medications had become increasingly available and acceptable during the decade that preceded 9/11. The public had heard about the benefits of Prozac, and about its hazards as well, and had voted for the former. Unfortunately, it took approximately 2 years for researchers to realize that the mental health-mediated debriefings meant to prevent PTSD had actually increased its occurrence. In contrast, those who were treated with medications that decreased physiological reactions fared better.

We might hypothesize that astute politicians and public health officials had an even greater goal when they choose to promote mental health treatment rather than religion to assuage stress. City bureaucrats, who are acutely aware of the intense role that religion plays in politics at any given time, may have realized that an excessive emphasis on religion could polarize people even more at this crucial time. Such

polarization could create serious social unrest, catalyze riots, trigger interreligious rifts, and provoke other unpredictable reactions. Perhaps someone realized that it was socially safer, if not clinically wiser, to shift the focus to universal psychiatric approaches and to deflect attention from particularistic religious relief.

Even though religion was not recommended as a first-line antidote for post-9/11 stress, we would be hard-pressed to deny the daunting impact of religion-related stress. Admittedly, we would be equally hard-pressed to find complete and accurate date on this topic at this early date, and so we must make do with speculation and hypotheses, seeded by anecdotal reports from the *New York Times* and garnered from personal clinical experiences.

Muslims and Mental Health after 9/11

Most obviously, we expect that Muslims would be affected most dramatically by the events of 9/11, not just because 9/11 was perpetrated by Muslims who acted against other Americans in the name of their faith but also because of the repercussions against Muslims (and against those who were mistaken for Muslims). At this point, it is important to point out that many different people from many different countries and from many different ethnic groups embrace Islam around the world and that the choices made by the suicide bombers were the choices made by but a few members of a multimillion-member religion. Yet all American-based Muslims were forced to face scrutiny by the American public. Many were subjected to police searches, involuntary incarceration, or automatic deportations, all of which caused stress.

Once the United States declared war on Iraq, American Muslims faced further stress, regardless of their personal political persuasions or religion affiliations. There were some similarities to the indignities endured by Japanese Americans during World War II, when their allegiance to the United States was suspect after Japan bombed Pearl Harbor, to the point that even U.S. citizens were moved out of their homes and resettled in closely guarded camps. Because Muslim Americans are sometimes (but certainly not always) identifiable through physical appearance or religious garb or Mideastern surnames, almost any Muslim can become an easy target of suspicion, discrimination, or outright abuse. Guilt through association can cause an assumption of guilt.

People who are in the United States on a student visa or a J-1 visa (as many physicians in training are) also carry legal documents that identify their country of origin. It would not be surprising if stress-related disorders increased among Muslims after 9/11, but it is not known with certainty if this actually occurred. The U.S. medical literature is strikingly devoid of studies on this subject at the time of this writing. In fact, it is doubtful that the full extent of the medical or mental health effects on Muslims will ever be known with certainty. It is likely that many Muslims, especially those who are working or studying in the United States on revokable visas, will avoid treatment altogether for fear of calling attention to themselves and risking being reported to the authorities or of appearing to be less qualified to compete in the work or educational arenas that influence their visas. It is probable that Muslims whose immigration papers are not complete or perfect will follow the paths taken by other illegal immigrants in the United States and will also be left uncounted and untreated by the official U.S. health-care system.

Furthermore, it is not clear whether psychiatric symptoms alone, such as stress, anxiety, or insomnia, could accurately measure the distress experienced by this very diverse religious group, whose members come from a wide variety of cultures, educational backgrounds, and countries of origin, and who speak many different languages. Some cultural groups do not express distress in psychological terms but prefer physical descriptions instead, sometimes complaining of dizziness, sleep problems, or vague aches and pains. Even people who are aware of psychiatric symptoms per se are still not immune to medical illnesses that increase in the presence of psychological or social stress. Cardiovascular and cerebrovascular events are notorious for their responsiveness to stress, as are some gastrointestinal diseases, autoimmune disorders, and even neurological symptoms.

In many instances, increases in drug and alcohol use are more accurate and immediate indicators of stress than are purely psychological symptoms and must be studied with as much diligence as medical or psychiatric symptoms. Alcohol consumption in New York City skyrocketed after 9/11, especially among women, as both medical statistics and bar revenues attested. But it is important to be aware of the baseline alcohol and drug consumption of any particular culture before determining if there has been a significant increase in use. For instance, Islam (apart from renegade Sufi mystics) generally prohibits alcohol. Therefore, even a modest increase in alcohol consumption in Muslims after 9/11 carries very serious clinical implications and cannot be meaningfully compared to increases that occurred in religious or ethnic groups that do not outlaw alcohol. On the

other hand, some Islamic cultures permit public marihuana use or even the social consumption of opium-based tea, although these activities are outlawed in the United States.

Although I cannot make claims to any population-based data on the use of psychiatric services by Muslims at this time, I can comment on recent trends in my own private practice in downtown Manhattan. Two distinctive gender-based trends emerged among Muslims.

First, there was a sudden surge in foreign-born Muslim males living in the United States on student visas who requested outpatient treatment for anxiety, insomnia, concentration problems, and vague perceptions of persecution by peers and co-workers. In some instances, patients needed an official return-to-work clearance, which they thought could be completed in a single visit. At other times, people who had long-standing behavior issues that could easily have been identified by college advisors much earlier were suddenly sent for psychiatric evaluations.

In each instance, these men were asked specifically if they thought that their new onset of psychiatric problems resulted from religious or cultural issues. If relevant, each was asked if he felt that the school or employer's request for psychiatric clearance was prompted by increased security concerns that were specific to Muslims. In each instance, the prospective patient denied any such association, but, in one case, the patient refused to fill out a form that inquired about native language, religion, and culture and made a point of stating that he did not consider religion to be relevant. When told that many patients, particularly college students, do not complete that section, he continued to insist that doctors do not need to know his religion.

Denial is not unusual in psychiatric practice, and a patient's repeated denial of the importance of an otherwise important issue can lead to a fertile investigation of psychological conflicts. Such denial usually takes place within the psychiatric interview itself and is addressed, often quite productively, during the therapy session. What was unusual in this situation was the prospective patient's repeated refusal to fill out standard medical history forms prior to the appointment.

Occasionally, the private outpatient psychiatrist encounters a paranoid patient who refuses to complete paperwork at the time of the first interview, but, more commonly, someone who is that paranoid avoids an appointment altogether and is brought to an institutional setting involuntarily. Moreover, pure paranoid patients typically show other observable signs of paranoia; their eyes scan the room and they sit near the door, if they sit down at all. They may refuse to enter enclosed spaces in the first place (which comes as welcome relief to the psychiatrist, who is equally ill-at-ease in being with a potentially volatile paranoid person).

Yet, in the situations described here, the Muslim men did not exhibit other observable signs of paranoia. Their previous social functioning prior to the onset of their psychiatric symptoms was definitely not suggestive of serious preexisting psychiatric illness. In each instance, these men were well-educated and performed well in the past, making it unlikely that their circumscribed suspicions about filling out forms was part and parcel of more global paranoia and appeared to be based on legitimate fears about incarceration or deportation or the loss of his job or student status or visas required for school or work, even though each denied this possibility. This para-paranoid behavior can be compared to the guardedness and evasiveness seen in some African American men who have had unpleasant experiences with legal authorities in the past.

Second, I noted an even greater increase in the number of Muslim American women seeking psychiatric treatment, albeit for very different reasons. These women typically sought treatment because of stress brought on by family conflicts that arose as a result of their defiance of the gender-role restrictions traditional in Muslim families. In some cases, romantic involvements with non-Muslim men had already caused serious family rifts and threatened the patient with the loss of both family and financial support. It is difficult, and also probably inaccurate, to attribute these situations solely to the aftershocks of 9/11. It is much more likely that Americanization in general was responsible for rattling the religious traditions of these new immigrants, just as Americanization and the increased social choices afforded by U.S. society have shaken the foundations of nearly every ethnic group that arrived on U.S. shores. Still, I wonder if the Muslim-mediated suicide bombings of 9/11 shook the faith of at least some Muslim women and gave them tacit permission to question the overall correctness of preexisting, previously unquestionable cultural and religious traditions. Alternatively, perhaps the widespread publicity about psychiatric treatment that came in the wake of 9/11 also made these women aware of the availability of secular psychiatric treatment for personal stress.

Jewish Issues Associated with 9/11 Stress

Last but not least, let me turn to specific stresses experienced by some Jews after 9/11. Although there are no studies to date that show that New York City Jews were

at increased risk for post-9/11 PTSD, there are reasons why Jews might experience different stresses than other religious groups. For one thing, some Muslim extremists said that they had bombed the United States specifically because of its support for Zionists in Israel. Moreover, some semiofficial spokespeople stated that they had targeted New York City in particular not just because of its centrality to the financial and communications industries but because of its disproportionately high Jewish population. Such extremists link all Jews to extremist Zionists, regardless of their personal political persuasions.

Jews might feel themselves targeted by the Muslim extremists, but they might also have residual PTSD because of personal or cultural connections to the Holocaust, when 6 million Jews perished at the hands of the Nazis. Those who smelled the burning bodies in the WTC – as did anyone who lived downtown, often for months afterward – typically compared the scene to the crematoria where Jewish bodies were burned en mass after being gassed in places such as Auschwitz. Such scents were described in detail in accounts by Holocaust survivors. Ironically, 2000 people perished in a single day at WTC on 9/11. That very same number – 2000 – was gassed every single day at Auschwitz when the ovens were at their peak performance. This strange synchronicity did not go unnoticed by those who knew the details of the Holocaust, and it served only to highlight the horrors of the Holocaust and to reopen old wounds that refused to heal.

Some Jews were pretraumatized because friends or relatives had been hurt or killed during recurrent terrorist attacks in Israel. Many felt personally threatened when they saw that the same terror could take place on supposedly safe U.S. soil. Among Jews who were familiar with Jewish history, 9/11 was compared to Kristallnacht, the day that marked the beginning of the Holocaust, when Jewish shops were shattered by Nazi storm troopers in a foreshadowing of the horror that was to come a few years later. The fact that this was the second, not the first, attack on the WTC was used to fortify the comparisons with Kristallnacht.

It is curious that a disproportionate number of Jews lived in the expansive Grand Street Co-ops in lower Manhattan. It is also curious that New York University, a university often identified as having a high proportion of Jewish students, dots downtown Manhattan. With its sprawling campuses and dormitories, New York University spans the area from the Seaport to Fifteenth Street, with some dorm rooms overlooking Ground Zero. It is possible that the presence of these ethnic and religious populations skewed statistics and influenced the high rate of PTSD among those who lived near Ground Zero. A more detailed demographic study of post-9/11 stress syndromes could clarify these possibilities.

Concerns of Christians and Others after 9/11

Is worth noting that members of some religions and ethnic groups became Muslims by proxy and were subjected to tauntings and even beatings because they were mistaken for Muslims. Turban-wearing Sikhs and Jains from the Indian subcontinent were targeted by aggressive youths in the days after the bombings and some were seriously injured. It is conceivable that people who practice many other religions but who appear to be of Mideastern or Muslim origin realized that they were at risk and so have suffered increased stress since 9/11.

It is likely that even all-American Christians also suffered from religion-related stress after 9/11, simply because most Americans like to think of themselves as religiously tolerant and become uneasy when religious prejudices are provoked, as they were by the events of 9/11. Let us not forget that the United States was founded for the expressed purpose of freedom of religion. When religion-inspired acts such as the terrorist bombings and hijackings of 9/11 threaten to make folly of this basic American freedom, everyone, regardless of his or her religion, feels stress.

In summary, Although there is much that remains uncertain about the overall impact of 9/11, both with respect to religion and stress, and to its many other ramifications, some things are certain. For one thing, 9/11 brought renewed attention to religion and its role in both promoting and protecting against stress. Unlike the Oklahoma bombing, which was strictly secular, this mass murder was intimately linked to religion and always will be, regardless of what is revealed about the individual bombers' personal psychologies in the future. This sad event serves as a reminder that the age-old association between stress and religion cannot be put on the shelf and relegated to the past; it is very much a part of the present. The subject of stress and religion is as relevant to the twenty-first century as it was to the first century CE.

Further Reading

Dervic, K., Oquendo, M. A. and Grunebaum, M. F. (2004). Religious affiliation and suicide attempt. *American Journal of Psychiatry* **161**, 2303–2308.

Durkheim, E. (1951). *Suicide*. Spaulding, J. A. & Simpson, G. (trans.). New York: Free Press.

Falsetti, S. A., Resnick, P. A. and Davis, J. L. (2003). Changes in religious belief following trauma. *Journal of Trauma and Stress* **16**, 391–398.

Galea, S., Vlahov, D. and Resnick, H. (2003). Trends of probably post-traumatic stress disorder in New York City after the September 11 terrorist attacks. *American Journal of Epidemiology* **158**, 514–524.

Koenig, H. G., McCullough, M. E. and Larson, D. B. (2001). *Handbook of religion and health*. New York: Oxford University Press.

Schuster, M. A., Stein, B. D. and Jaycox, L. (2001). A national survey of stress reactions after the September 11, 2001 terrorist attacks. *New England Journal of Medicine* **345**, 1507–1512.

Shalev, A. Y. (2004). Further lessons from 9/11: does stress equal trauma? *Psychiatry* **67**, 174–177.

Wolinsky, F. D., Wyrich, K. W. and Kroenke, K. (2003). 9–11, personal stress, mental health, and sense of control among older adults. *Journal of Gerontology, Psychology, and Social Sciences* **58**, S146–150.

Remodelling of Neuronal Networks by Stress

E Fuchs
German Primate Center and Medical School, University of Göttingen, Göttingen, Germany
G Flügge
German Primate Center, Göttingen, Germany

© 2007 Elsevier Inc. All rights reserved.

Early Ideas about the Stress Response
Stress Circuits
Activation of Brain Systems
Structural Plasticity of the Adult Brain
Remodeling of Brain Cells by Stress
Stress Suppresses Neurogenesis in the Adult Dentate Gyrus
Plasticity of Astrocytes
Plastic Changes of Brain Cells Are Reflected by Alterations in Gene Expression
Structural Changes Are Reversible
Conclusions

Glossary

BrdU (5-Bromo-2′-deoxyuridine)	A thymidine analog that is incorporated into the DNA of a dividing progenitor cell during the S phase. Fully differentiated neurons do not divide and cannot integrate the label. Depending on the survival time after application, BrdU is a marker of proliferating cells (short survival time) and their progeny (longer survival time). BrdU labeling is a nonisotopic immunocytochemical method.

Early Ideas about the Stress Response

When Selye formulated his stress theory approximately 70 years ago, stress was thought to have a merely endocrine character, and noxious stimuli of physical or chemical nature were the primary stressors discussed. However, subsequent research demonstrated that stress-induced activation of the endocrine system leads to increased release of glucocorticoids (cortisol in primates including humans, corticosterone in rodents) from the adrenal cortex and that these peripheral stress hormones also profoundly influence brain function. Moreover, stress also increases the activity of the sympathetic nervous system to enhance release of norepinephrine from peripheral nerve endings. In addition, stress modulates the activities of, for example, central nervous monoaminergic circuits. These have diverse effects on many brain cells. Great progress has been made in the understanding of the biology of stress reactions and the neurobiological alterations that occur as a consequence of exposure to challenging events in the environment.

Stress Circuits

One of the most intensively investigated stress circuits is the limbic-hypothalamic-pituitary-adrenocortical (LHPA) system, which determines the magnitude and duration of hormonal, neural, and behavioral reactions to stressful experiences. The LHPA system combines both brain and endocrine components and can be regarded as a classic neuroendocrine circuit that integrates cognitive, emotional, neuroendocrine, and autonomic inputs. It consists of the primary brain regions that process emotional information, the limbic system and the hypothalamus, as well as pituitary and adrenal glands as endocrine organs. In concert with other neurohormonal actions such as stimulation of the adrenomedullary system (release of adrenaline from the adrenal medulla), LHPA activation leads to an increased release of corticosteroid hormones from the adrenal cortex into the bloodstream. These

steroid hormones are potent modulators of cell physiology, mood, and behavior and serve both to alert the organism to an environmental or physiological challenge and to defend its homeostasis.

Within the LHPA system, the hippocampus plays a central role. This brain structure has a high density of corticosteroid receptors and modulates glucocorticoid release from the adrenal cortex via a negative feedback loop exerting inhibitory effects on the LHPA axis. In addition to its intimate involvement in neuroendocrine regulation, the hippocampus plays an important role in emotional processing and in key aspects of learning and memory. While short-term activation of the LHPA system is essential for vital functions, constant hyperactivity of the circuit results in chronically elevated glucocorticoids, a maladaptive and potentially neuropathological condition. As demonstrated in animal studies, chronic stress leads to an impairment of the negative feedback control mechanism of corticosteroid release from the adrenal cortex, partly because of a reduced expression of hippocampal corticoid receptors.

Activation of Brain Systems

Stress does not have a global effect on all brain areas. Rather, depending on the type of stressor, different neuronal circuits are activated. Limbic brain regions comprising, for example, the hippocampus and amygdala as well as the prefrontal cortex are sensitive to stressors such as restraint or anxiogenic external stimuli, for example, a novel environment. Common to these stressors is the fact that they require, prior to initiation or inhibition of the stress response, a central nervous processing of information coming from different sensory organs. This information processing involves limbic and cortical regions, and its outcome depends on previous experience. In contrast, physiological threats such as exposure to ether result in a direct activation of efferent visceral pathways, in part mediated by the paraventricular nucleus of the hypothalamus. In this case, the rapid activation of brain stem and hypothalamus circumvents cognitive-emotional processing via the higher brain regions.

The intraneuronal communication that finally leads to the stress reaction is brought about by neurotransmitter systems. Rapid neuronal responses to stress are elicited by, for example, glutamate (an excitatory neurotransmitter) and GABA (γ-aminobutyric acid, an inhibitory neurotransmitter). These modulators act on specific receptors and ion channels in neuronal membranes to directly control neuronal firing. This fast neurotransmission is modulated by slow-acting neurotransmitters such as the monoamines (for example, the catecholamines dopamine, norepinephrine, and epinephrine and the indolamine serotonin). Activation of brain monoamine systems is a major component of the stress response, and turnover of monoamines is increased during stressful experiences. Stress-induced alterations in the monoamine system are regarded as the basis for stress-related behavior and are of special interest because the hyperactivity of catecholaminergic and serotonergic neurons that occurs during stress may induce an imbalance between the different neurotransmitter systems. This imbalance may contribute to psychopathologies such as major depression.

Another example of substances that modulate neuronal activity and are involved in mediation of stress responses are neuropeptides, for example corticotropin-releasing hormone (CRH). This is thought to play a major role in the stress response. Neuropeptides as well as monoamines exert their actions via specific receptors located in the membranes of neurons. Some of these receptors are also found in glia cells. Many of these receptors belong to the family of G-protein-coupled receptors (GPCRs) that are linked to intracellular second messenger systems via guanosine 5′-triphosphate (GTP)-binding proteins. The intracellular signal transduction pathways that are regulated by GPCRs bring about many of the short- and long-term effects of stress on neuronal activity and regulate most aspects of neuronal functioning, including metabolism and gene expression. Thus, intracellular signal transduction systems play a key role with respect to stress-mediated regulation of neuronal functioning and structure, and, on the cellular level, they represent the first step of a cascade of events in the brain that leads to changes in neural cells, including alterations in gene transcription and morphology of neurons. Adverse effects of stress are supposed to result partly from (1) sustained activation or inhibition of certain neurotransmitter systems and intracellular pathways and/or (2) chronic exposure to neuroactive substances such as the glucocorticoids. These steroid hormones interact with intracellular receptors, forming complexes that bind to specific sites on DNA to regulate transcription.

Structural Plasticity of the Adult Brain

Approximately a century ago, it was finally established that the brain consists of neurons and glia (neural cells) rather than being a continuous syncytium. Thereafter, a rather static view of the brain prevailed. It was thought that electrical and chemical information was processed by fixed neuronal circuits. Ramon ý Cajal stated that “in adult centers the nerve paths are something fixed, ended, immutable. Everything may die, nothing may be regenerated” (Cajal

1928, p. 750). However, in recent years, this view has been gradually revised on the basis of studies demonstrating that neuronal circuits and connections between cells in the brain are subject to lifelong modifications and reorganizations. An impressive example of the consequences of alterations in sensory inputs and changes in circulating hormones comes from research using ground squirrels. In these hibernating animals, synaptic contacts of mossy fibers with CA3 hippocampal pyramidal neurons are altered in many aspects of their structure during different stages of the torpor–activity cycle. This clearly demonstrates the capacity of the adult brain to adjust its neuronal circuits to altered hormonal and/or sensory input. Based on such findings, it was hypothesized that similar structural modifications may also occur during stressful life events when hormonal levels are altered (for example, elevated glucocorticoids, reduced gonadal steroids) and when sensory input is changed. It has transpired that this view is accurate.

This ability of the brain to perform functionally relevant adaptations following various challenges is called plasticity. Neuronal plasticity is absolutely mandatory for adequate functioning of an individual in a continuously changing environment. The dynamic processes of neuroplasticity are based on the capability of neurons to adapt to alterations in the internal and/or external environment by modifying specific structures and function; they do this by changing their entire cell surfaces, including the synapses, so that the locations of receptors and other molecules imbedded in the plasma membrane are reorganized. Adaptive or experience-dependent plasticity reveals itself in many forms, ranging from changes in gene expression to changes in neurotransmitter release and in behavior. Even under normal or undisturbed conditions, contacts between neurons are continuously replaced and renewed in the adult and differentiated brain. Enhanced axonal outgrowth and collateral sprouting on the presynaptic site may lead to the formation of new synapses, and existing contacts between neurons may be eliminated by terminal retrograde degeneration. The number of postsynaptic sites on a neuron can be increased or decreased by alterations in the size of its dendritic tree or the spine density on the dendrites.

Remodeling of Brain Cells by Stress

Probably the most thoroughly investigated stress-induced change in neuronal morphology is the regression of apical dendrites of pyramidal neurons, which was first demonstrated in the hippocampus. After this finding, dendritic remodeling of CA3 pyramidal neurons has been repeatedly documented after chronic stress as well as after corticosterone administration. Similar shortening of dendritic branches, although to a smaller extent, has been observed in granule cells of the dentate gyrus and in pyramidal cells of the hippocampal region CA1, both in chronically stressed and in corticosterone-treated rats. It is assumed that, as a result of the reduced surface area of the neurons, synaptic contacts are also diminished. Indeed, a significant loss of synapses on CA3 pyramidal cells, as well as profound changes in the morphology of the afferent mossy fibers that terminate on these neurons, was detected in chronically stressed or corticosterone-treated animals. Interestingly, even a brief social defeat stress with a long time delay thereafter can significantly reduce the length of apical dendrites but increase the length and complexity of the primary dendrites at the basis of the pyramidal neurons. These data suggest that a brief social conflict is sufficient to drive a dynamic reorganization of neuronal networks with site-selective elimination as well as *de novo* growth of dendritic branches, changes that persist for several weeks after the acute stress exposure.

Contrary to conclusions from initial studies proposing an impact of chronic stress on the number of pyramidal neurons in the hippocampus, recent stereological investigations failed to detect a statistically significant loss of these neurons following prolonged hypercortisolism resulting from stress exposure, corticosteroid administration, or aging. However, it should be noted that the number of hippocampal interneurons or of neurons in the hilus were not specifically examined in these studies. Recently, evidence was provided that long-term psychosocial stress reduces the number of parvalbumin-immunoreactive neurons that are regarded as GABAergic interneurons. This stress effect was observed in the dentate gyrus and in hippocampal region CA3, whereas subfield CA1 was not affected. Thus, the possibility that long-term stress may induce loss of interneurons without affecting the number of principal cells in the hippocampus cannot be ruled out.

Chronic stress alters dendritic morphology not only in the hippocampus but also in the medial prefrontal cortex (mPFC), which modulates various higher cognitive and emotional functions and contributes to regulation of LHPA system activity. Recent reports demonstrated that 3 weeks of either daily restraint stress or daily corticosterone injections led to changes in pyramidal neurons of the mPFC. Up to 20–35% retraction of the distal dendritic branches was seen, together with a significant (16%) decrease in apical dendritic spine density. Interestingly, even repeated vehicle injections resulted in similar, although less pronounced, changes, indicating that with respect to morphology of neurons, the mPFC is even more sensitive to external influences than the hippocampus.

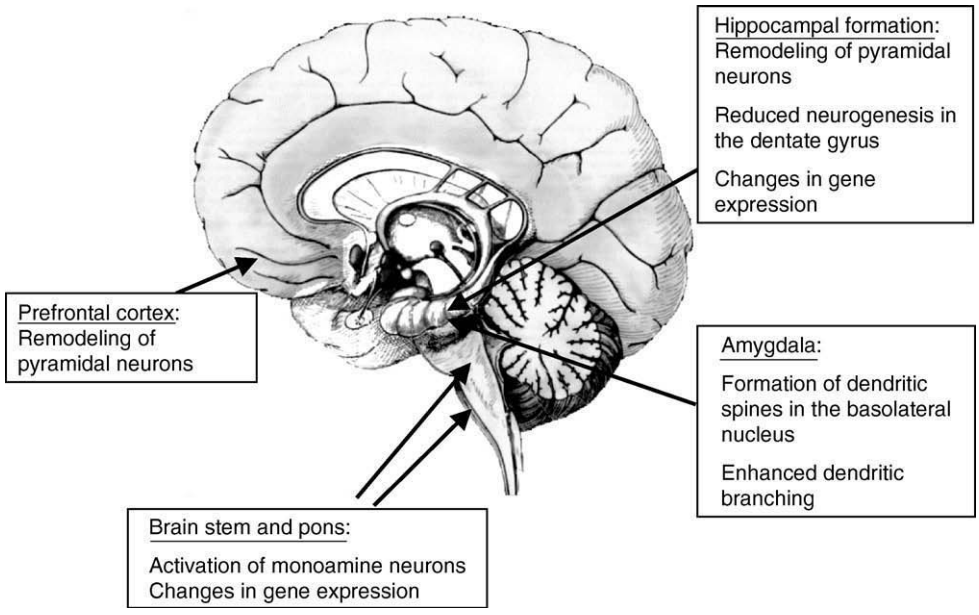


Figure 1 Examples of stress-induced changes in the brain. Hippocampal formation: remodeling of pyramidal neurons characterized by shrinkage of dendrites of pyramidal neurons in region CA3; changes in mRNA and protein reflect that stress regulates expression of genes involved in neuronal differentiation. Prefrontal cortex: remodeling of pyramidal neurons. Amygdala: in the basolateral nucleus, formation of new spines on the dendrites of spiny neurons. Brain stem and pons: activation of monoamine neurons, e.g., those that produce norepinephrine; changes in gene expression.

The amygdala plays a crucial role in regulating fear and anxiety and is essential for the formation of affective memories. Increasing evidence supports a critical role of this brain region in activating the LHPA system. A recent study investigated the impact of chronic stress on the dendritic pattern of amygdaloid neurons. In contrast to findings in the hippocampus, spiny neurons in the basolateral complex of the amygdala exhibited enhanced dendritic branching in response to chronic stress. Examples of stress-induced changes in the brain are summarized in Figure 1.

Stress Suppresses Neurogenesis in the Adult Dentate Gyrus

For a long time, a central hypothesis in neuroscience was that in the mammalian brain, the production of neurons occurs only during development and stops before puberty, and new neurons cannot be formed in the adult brain. This widely held belief has been challenged in recent years by extensive evidence from many mammalian species, including nonhuman primates as well as humans, showing that even in the adult brain, certain areas retain the capability to generate new neurons. Adult spontaneous neurogenesis takes place only in certain regions such as the subgranular zone of the dentate gyrus in the hippocampal formation and the subventricular zone at the lateral ventricle. In the dentate gyrus, newly generated cells become incorporated into the granule cell layer and attain the morphological and biochemical characteristics of granule neurons. The neuronal nature of these cells is demonstrated by the formation of synapses on the cell bodies and dendrites, extension of axons into the CA3 region, and generation of action potentials.

These observations on neurogenesis were facilitated by the advent of a novel method for detecting cell proliferation and migration. Using the thymidine analog 5-bromo-2′-deoxyuridine (BrdU), which is incorporated into the DNA of dividing cells during the S phase of mitosis, proliferating cells and their progeny can be labeled. BrdU can be visualized by immunocytochemical techniques, and when used in combination with other specific cell markers, it allows determination of the phenotype of the newly generated cells. Using such methods, it has been demonstrated that the majority of newborn cells in the dentate gyrus differentiate into neurons. In other words, it has been shown that the BrdU-labeled cells express the neuronal marker neuron-specific nuclear protein. A smaller proportion of the newborn cells express markers for glial cells such as glial fibrillary acidic protein. Although the new neurons comprise a minuscule proportion of the total neuronal population, their continuous addition over the entire life span implies considerable structural changes within the neural network. The magnitude and ubiquity of adult neurogenesis across vertebrates suggest that it is

functionally significant and not merely a vestige of development.

Stress is a very potent suppressive factor of adult neurogenesis, although a large number of environmental and endogenous parameters also have an impact on the formation of new neurons in the adult brain. The suppression of cell turnover by stress predicts that the age of the cell population, the connectivity of the neurons, and the resulting properties of neuronal circuits in a stressed individual might be substantially different from control situations, with potentially important functional consequences. Several different types of stressful experiences, such as exposure to the odor of a predator, social subordination, or resident–intruder stress, have been shown to inhibit neurogenesis in the dentate gyrus in several species, including nonhuman primates. Furthermore, it appears that the susceptibility of neurogenesis to stress is age dependent in that effects of stress on cell proliferation in the dentate gyrus are more pronounced in old animals than in young ones.

It has become increasingly evident that the antecedents of many illnesses begin in fetal life and, furthermore, that prenatal conditions can have a significant impact on either health or disease in the postpartum period. In a recent study that investigated whether prenatal stress can alter neural, hormonal, and behavioral status in nonhuman primates, pregnant rhesus monkeys were acutely stressed on a daily basis for a quarter of their 24-week gestation using an acoustic startle protocol. At 2 to 3 years of age, hippocampal volume, cytogenesis in the dentate gyrus, and cortisol levels were evaluated in the offspring from stress and control pregnancies. Prenatal stress reduced hippocampal volume and inhibited neurogenesis in the dentate gyrus. These changes were associated with higher cortisol levels in the blood of the monkey offspring, lower levels of exploration, and higher levels of motor behavior. These findings indicate that also in primates, a prenatal environment can alter behavior, may deregulate neuroendocrine systems, and affects the hippocampal structure in a persistent manner.

Plasticity of Astrocytes

The most abundant type of cell within the central nervous system is the glia. In the adult mammalian brain, there are 10 to 50 times more glia cells than neurons. Astrocytes are the most prominent type of glia cells, accounting for about one-third of brain mass. Recent studies have revealed that, in addition to their housekeeping functions, astrocytes are dynamic regulators of synaptogenesis and synaptic strength. These cells control neuronal production, for example, in the adult dentate gyrus. Thus, abnormalities in glial functioning are likely to contribute to the impairment of structural plasticity in the brain. In line with these ideas, it was recently demonstrated that long-term stress significantly decreased both the number and the volume of astroglial somata. In addition, whether antidepressant treatment with the selective serotonin reuptake inhibitor fluoxetine could afford protection from these stress effects was examined. Indeed, fluoxetine prevented a stress-induced decrease in the number of astrocytes but did not counteract shrinkage of the astroglial somata.

Morphological changes in astrocytes most probably have functional significance with respect to neuron–glia interaction, and because this interaction supports neuronal functioning, interneuronal communication will also be affected by the stress-induced changes. The reduced number or weakened activity of astrocytes could change levels of extracellular glutamate, thus leading to high concentrations of this excitatory neurotransmitter that may have excitotoxic effects. An upregulation of the glial glutamate transporter (GLT-1) in the hippocampus has been reported after chronic stress, and it has been suggested that this might be a compensatory mechanism to control for increased extracellular glutamate concentrations caused by stress. Interestingly, a treatment with the antidepressant tianeptine can block the stress-induced upregulation of GLT-1, and tianeptine also reverses stress-induced dendritic remodeling in CA3. Therefore, the modulation of glial GLT-1 expression in hippocampal region CA3 can be considered a regional neurochemical correlate of dendritic remodeling. These changes in the structure and molecular plasticity of astroglia in response to stress and antidepressant treatment support the notion that glial changes may contribute to the pathophysiology of stress-related disorders such as major depression as well as to the cellular actions of antidepressants.

Plastic Changes of Brain Cells Are Reflected by Alterations in Gene Expression

The stress-induced changes in neuronal structure imply that brain cells adjust their biosynthetic pathways to the new requirements. For example, concomitant with retraction of dendrites, there is probably a reduced need for membrane proteins. Related adaptational processes take place partly on the level of gene transcription. Because various methods, namely, cDNA microarrays, serial analysis of gene expression, and subtractive hybridization, are currently used to

identify genes that are differentially regulated by stress, such genes still appear rather heterogeneous. However, it is already clear that some observations from gene transcription studies are consistent with what has been found with other methods.

In hippocampal formation, expression of several genes known to be involved in neuronal differentiation was found to be downregulated by chronic social stress. Such genes included those encoding the membrane glycoprotein M6a, the CDC-like kinase 1, and a gene encoding a distinct subunit of certain G-proteins, GNAQ. All these genes are known to be involved in neurite outgrowth and neuronal differentiation, thus supporting the view that alterations in neuronal morphology and/or formation of neurons are primary effects of stress, at least in hippocampal formation. Furthermore, expression of neural cell adhesion molecule (NCAM) was found to be downregulated after chronic restraint stress in the hippocampus. NCAM is known to regulate neurite outgrowth and target recognition in the developing nervous system by mediating cell adhesion and regulating signal transduction.

Neuronal plasticity is accompanied by dynamic changes in elements of the cytoskeleton. Alpha-tubulin, which is the major component of microtubules, can be posttranslationally modified, and tyrosinated and acetylated alpha-tubulin are considered markers of dynamically changing and stable forms of microtubules, respectively. Restraint stress for 4 days decreased the expression of tyrosinated alpha-tubulin and increased expression of the acetylated form in the hippocampus. This is in line with the view that stress alters the morphology of pyramidal neurons.

Plastic changes are not restricted to forebrain areas. In a brain stem region, the rat dorsal raphe nucleus, which contains a large number of serotonergic neurons innervating the forebrain, long-term social stress increased the expression of genes involved in regulation of neurotransmitter release (synaptosomal-associated protein 25 and synaptic vesicle glycoprotein 2b). These data support the view that stress increases activity of neurons in the dorsal raphe nucleus.

Data from mRNA or protein expression studies sometimes appear contradictory. However, when interpreting such data, in which brain region, on which cellular level (neurons versus glia), and in which cellular compartment (neuronal cell body versus area of the dendrites versus axon terminal areas) the stress-induced changes were detected should be considered. Finally, besides the sex and the age of the investigated species, both the type and the duration of the stressor play a crucial role with respect to effects of stress on gene expression.

Structural Changes Are Reversible

As shown in animal studies, apical dendrites of hippocampal pyramid neurons that were shortened by restraint stress or glucocorticoid exposure regained their normal length within 3 weeks after treatment. Several *in vivo* magnetic resonance imaging studies in humans suggest a correlation between hippocampal volume reduction and cumulative glucocorticoid exposure, although exceptions have also been reported. In Cushing's patients who suffer from hypercortisolism, smaller hippocampal volumes returned to normal after successful surgery that led to normal glucocorticoid levels. Thus, up to a certain point, structural changes may be at least partially reversible and may permit restoration of normal functioning. However, it appears that the appropriate intervention should take place within a certain time before the changes become irreversible.

The mechanisms that are responsible for hippocampal volume loss have not yet been identified. Massive neuronal loss following exposure to repeated episodes of hypercortisolemia can be excluded, as no major cell loss was apparent in experimental animals as well as in human postmortem brain tissue. Neurogenesis has been regarded as another mechanism that might affect the volume of the hippocampus. However, hippocampal neurogenesis adds relatively few neurons per day, and those only to the granule cell layer of the dentate gyrus, where an equivalent number of neurons also die. Alternately, hippocampal volume loss might be due to alterations in amounts of thedendritic, axonal, and synaptic components or to changes in levels of glia. Furthermore, stress may induce a shift in fluid balance between the ventricles and brain tissue. This assumption is supported by numerous clinical studies reporting lower volumes of different brain structures in conjunction with enlarged ventricles in patients with affective disorders. However, it remains to be determined whether changes in fluid content cause the aforementioned neuroplastic changes in the neuronal network or just accompany them.

Conclusions

The diverse forms of stress-induced changes in neural cells are of particular interest. They show that even the adult and differentiated brain is a plastic organ. This lifelong plasticity is a prerequisite for the brain to adapt to environmental changes. Plasticity of the neuronal networks is the basis for learning and memory. On the other hand, extensive neuroplastic changes induced by stress may also lead to imbalances in central nervous neural networks. A promising aspect with regard to therapy is that many of these

processes appear to be reversible. Reversibility of structural as well as functional plasticity has already been demonstrated, for example, in response to pharmacological treatments. Thus, progress in our understanding of neural plasticity has profound implications for the treatment of a number of neurodegenerative and psychiatric disorders.

See Also the Following Articles

Acute Trauma Response; Corticosteroid Receptors; Glucocorticoid Effects on Memory: the Positive and Negative; Glucocorticoids – Adverse Effects on the Nervous System; Memory and Stress; Maternal Deprivation; Peptides.

Further Reading

Abumaria, N., Rygula, R., Havemann-Reinecke, U., et al. (2006). Identification of genes regulated by chronic social stress in the rat dorsal raphe nucleus. *Cellular and Molecular Neurobiology* **26**, 145–162.

Alfonso, J., Frasch, A. C. and Flügge, G. (2005). Chronic stress, depression and antidepressants: effects on gene transcription in the hippocampus. *Reviews in Neuroscience* **16**, 43–56.

Cajal, S. R. (1928). *Degeneration and regeneration of the nervous system*. London: Oxford University Press.

Czeh, B., Simon, M., van der Hart, M. G., et al. (2005). Chronic stress decreases the number of parvalbumin-immunoreactive interneurons in the hippocampus: prevention by treatment with a substance P receptor (NK1) antagonist. *Neuropsychopharmacology* **30**, 67–79.

Czéh, B., Simon, M., Schmelting, B., Hiemke, C. and Fuchs, E. (2005). Astroglial plasticity in the hippocampus after chronic psychosocial stress and concomitant fluoxetine treatment. *Neuropsychopharmacology,* Dec. 14 [Epub ahead of print].

de Kloet, E. R., Vreugdenhil, E., Oitzl, M. S. and Joëls, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* **19**, 269–301.

Flügge, G., van Kampen, M. and Mijnster, M. J. (2004). Perturbations in brain monoamine systems during stress. *Cell and Tissue Research* **315**, 1–14.

Fuchs, E. and Czéh, B. (2005). Adult neurogenesis in rodents and primates: functional implications. In: Steckler, T., Kalin, N. H. & Reul, H. (eds.) *Handbook of stress and the brain Part 1: The neurobiology of stress. Techniques in the behavioral and neural sciences* (Vol. 15), pp. 711–727. Amsterdam: Elsevier.

Fuchs, E., Czéh, B., Kole, M. H. P., Michaelis, T. and Lucassen, P. J. (2004). Alterations of neuroplasticity in depression: the hippocampus and beyond. *European Neuropharmacology* **14**, 481–490.

Herman, J. P. and Cullinan, W. E. (1997). Neurocircuitry of stress: central control of the hypothalamo-pituitary-adrenocortical axis. *Trends in Neuroscience* **20**, 78–84.

Kempermann, G. (2002). Regulation of adult hippocampal neurogenesis – implications for novel theories of major depression. *Bipolar Disorders* **4**, 17–33.

Mason, J. W. (1968). A review of psychoneuroendocrine research on the pituitary-adrenal cortical system. *Psychosomatic Medicine* **30**, 576–609.

McEwen, B. S. (1999). Stress and hippocampal plasticity. *Annual Review of Neuroscience* **20**, 49–70.

Mitra, R., Jadhav, S., Bruce, S., et al. (2005). Stress duration modulates the spatiotemporal patterns of spine formation in the basolateral amygdala. *Proceedings of the National Academy of Sciences USA* **102**, 9371–9376.

Popov, V. I. and Bocharova, L. S. (1992). Hibernation-induced structural changes in synaptic contacts between mossy fibres and hippocampal pyramidal neurons. *Neuroscience* **48**, 53–62.

Reul, J. M., Sutanto, W., van Eekelen, J. A., Rothuizen, J. and de Kloet, E. R. (1990). Central action of adrenal steroids during stress and adaptation. *Advances in Experimental Medicine and Biology* **274**, 243–256.

Sapolsky, R. M. (1999). Glucocorticoids, stress, and their adverse neurological effects: relevance to aging. *Experimental Gerontology* **34**, 721–732.

Selye, H. (1936). A syndrome produced by diverse noxious agents. *Nature London* **138**, 32.

Stanford, S. C. (1993). Monoamines in response and adaptation to stress. In: Stanford, S. C. & Salmon, P. (eds.) *Stress. From synapse to syndrome*, pp. 281–331. London: Academic Press.

Watanabe, Y., Gould, E. and McEwen, B. S. (1992). Stress induces atrophy of apical dendrites of hippocampal CA3 pyramidal neurons. *Brain Research* **588**, 341–345.

Wellman, C. L. (2001). Dendritic reorganization in pyramidal neurons in medial prefrontal cortex after chronic corticosterone administration. *Journal of Neurobiology* **49**, 245–253.

Renal and Adrenocortical Actions of Dopamine

B C Williams, Y-C Lo and S W Walker
University of Edinburgh, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 356–359, © 2000, Elsevier Inc.

Renal Actions of Dopamine
Adrenocortical Actions of Dopamine
Conclusions

Glossary

Aromatic L-amino acid decarboxylase (L-AAAD) (EC 4.1.1.28) — Previously known as hydroxytryptophan decarboxylase or dopa decarboxylase, L-AAAD catalyzes the decarboxylation of L-5-hydroxytryptophan and L-dopa (L-3,4-dihydroxyphenylalanine) to 5-hydroxytryptamine and dopamine, respectively. It requires pyridoxal phosphate as coenzyme and occurs most abundantly in the kidney, liver, gastrointestinal tract, adrenal gland, and brain.

Dopamine (3,4-dihydroxyphenylethylamine) — A biogenic amine, first identified as an important neurotransmitter in the central nervous system and later discovered to exert actions in peripheral tissues, particularly the cardiovascular system, the renal system, and the adrenal gland.

Dopamine receptors — Plasma membrane proteins that mediate the biological effects of dopamine through the activation of an intracellular second messenger system. Many different classes of dopamine receptors have been defined according to their pharmacological properties using selective agonists and antagonists and/or their molecular structure determined by molecular cloning techniques. Dopamine receptors in peripheral tissues are still distinguished from those in the central nervous system. Although they appear to be similar, they may not be identical, thus D_1 and D_2 denote dopamine type 1 and type 2 receptors, respectively, in the central nervous system, whereas DA_1 and DA_2 denote what are considered to be the equivalent (and perhaps related) receptors in peripheral tissues.

Natriuresis — The loss of sodium ions from circulating plasma into the urine through the action of endogenous chemicals or drugs on the sodium transport process occurring in the proximal and distal segments of the renal tubules.

Zona glomerulosa — A narrow layer of cells in the outer region of the adrenal gland, situated directly beneath the adrenal capsule, whose specific function is to secrete the steroid hormone aldosterone, which acts on mineralocorticoid receptors in the distal convoluted tubule of the kidney to cause sodium retention and potassium loss.

Renal Actions of Dopamine

Much of our current understanding of the renal actions of dopamine derives from the pioneering work of Leon Goldberg from 1964 onward, which stimulated subsequent research throughout the world.

Effects of Dopamine on the Renal Vasculature

Infusion of dopamine *in vivo* in several animal models leads to vasodilatation and to an increase in renal blood flow with a less marked rise in the glomerular filtration rate (GFR). The limited studies conducted in humans have confirmed these effects, and it appears that the major effect of dopamine is to increase selectively blood flow through the renal cortex rather than through the renal medulla, indicating that dopamine causes a redistribution of renal blood flow, which may reflect the relative distribution of dopamine receptors in the renal cortex versus the renal medulla vasculature. Dopamine receptors that mediate dopamine-induced renal arterial dilatation have been partially characterized using pharmacological agents. In the *ex vivo*-perfused rat kidney, the DA_2 receptor agonist bromocriptine causes vasodilatation, and this response is inhibited by metoclopramide, a DA receptor antagonist. In contrast, in renal microvessels from rabbits, the dopamine-induced relaxation of afferent and efferent glomerular arterioles could be mimicked by fenoldopam and SKF 87516, two specific DA_1 receptor antagonists. In humans, the infusion of fenoldopam causes renal vasodilatation, and in human renal artery segments, dopamine-induced relaxation can be blocked by (+)-sulpiride and SCH 23390, two DA_1 receptor antagonists. It therefore appears that both DA_1 and DA_2 receptors play a role in dopamine-induced vasodilatation and that their respective contribution may be species specific.

Effects of Dopamine on the Renal Tubule

In pharmacological doses, dopamine is a potent natriuretic agent in the kidney. The mechanisms for this natriuretic action appear to be multiple, involving increases in renal plasma flow and GFR, a redistribution of blood flow, and inhibition of the tubular transport of sodium. Dopamine infusion also increases urine cyclic AMP excretion in humans and this effect parallels the increased excretion of sodium into the renal tubules. There is now clear evidence for the involvement of DA_1 receptors in the natriuretic response to dopamine. Infusions of the DA_1 receptor antagonist SCH 23390 via the renal artery in the conscious uninephrectomized dog decrease urine flow rate and sodium excretion, and this effect is inhibited by the DA_1 receptor agonist fenoldopam. These *in vivo* studies have been supported by *in vitro* experiments, which have demonstrated specific binding sites for the DA_1 antagonist SCH 23390 in the proximal tubule of several species, and dose-dependent inhibition of the enzyme Na^+K^+-ATPase in the thick ascending limb of the loop of Henle by the DA_1 receptor agonist fenoldopam, which can be blocked by the DA_1 receptor antagonist SCH 23390. The role of DA_2 receptors in the natriuretic action of dopamine still requires further clarification, although these receptors have been identified clearly in the rat proximal tubule.

Effects of Dopamine on Renin Release

The effects of dopamine on renin release are variable, depending on the species and the experimental model. Intrarenal infusion of dopamine in the conscious dog causes a dose-dependent increase in renin release, which is attenuated by the DA_2 receptor antagonist sulpiride. Similarly, in both the isolated perfused rat kidney and dispersed rat renal cortical cell preparations, dopamine stimulates renin release via the activation of DA_2 receptors. In humans, the intravenous infusion of dopamine has been shown either to have no effect on plasma renin activity (PRA) or to increase PRA. In contrast, oral administration of the DA_2 receptor agonist bromocriptine, in humans, causes a decrease in PRA. It has been hypothesized that the stimulatory effects of dopamine on renin release may relate to a direct postsynaptic effect on the renal juxtaglomerular cells and that the inhibitory actions of dopamine on renin release may be presynaptic. Further studies are essential to clarify the actions of dopamine on renin release, especially in humans.

Intrarenal Sources of Dopamine

Renal plasma concentrations of dopamine are low compared to the concentrations of dopamine required to cause the effects described earlier. It is also apparent that the concentration of dopamine found in urine is vastly greater than what could be accounted for by filtered plasma dopamine. This led to the hypothesis that there were renal sources of dopamine. One hypothesis suggested that dopamine in the kidney was largely derived from the dopamine conjugates sulfate and glucuronide, which are formed in the adrenal medulla and deconjugated in the kidney by the renal enzymes arylsulfatase and β-D-glucuronidase, respectively, releasing free dopamine. Although dopamine conjugates may be one source of renal dopamine, they do not appear to be the major source, as sodium chloride loading in adrenalectomized rats still leads to a marked increase in urine dopamine. A second hypothesis suggested that dopaminergic nerves, which were shown by immunofluorescence studies to be located primarily at the glomerular vascular pole, could release dopamine for local actions on renal hemodynamics in the kidney. In support of this hypothesis, it was demonstrated that renal nerve stimulation in the rat kidney produced vasodilatation that could be inhibited by the DA_2 receptor antagonist sulpiride. It seems likely that dopaminergic nerves could play a role in the regulation of glomerular tuft blood flow and in the regulation of renin release from the juxtaglomerular cells in the afferent arteriole. It is less likely that this source of dopamine plays a role in the natriuretic action of dopamine in the renal tubules. The major source of renal dopamine, which is of specific relevance to the natriuretic effects of dopamine, is undoubtedly that produced by the action of the enzyme aromatic L-amino acid decarboxylase (L-AAAD) on circulating L-dopa. This enzyme is located in the renal tubules, with higher levels in the proximal as compared to the distal segments. An increase in sodium chloride intake causes an increase in renal dopamine and urine dopamine, which may reflect an increase in the uptake of L-dopa from plasma into the proximal tubule (which is a sodium-dependent process), an increase in L-AAAD activity, or both. When the specific inhibitor of L-AAAD, carbidopa, is administered to humans, urine sodium excretion falls as a consequence of the decrease in renal dopamine synthesis.

Pathophysiology of Renal Dopamine

Because the renal actions of dopamine (increased GFR, vasodilatation, and natriuresis) are beneficial for the maintenance of normal renal function, its possible role in several different pathophysiological states, where renal function may be compromised, has been investigated. In some patients with essential hypertension, urinary dopamine excretion is reduced, and an increased sodium intake produces a

paradoxical fall in urinary dopamine in comparison to normal subjects, where sodium loading leads to an increase in urinary dopamine. This led to the hypothesis that the increased retention of sodium in some forms of human hypertension may relate to a defect in the kidney to produce dopamine in order to excrete the excess sodium from the body. Similarly, patients with chronic renal failure show reduced dopamine excretion compared to normal subjects and an impaired renal vascular response to dopamine infusions, which may relate to a loss of dopamine receptors within the kidney.

Adrenocortical Actions of Dopamine

Dopamine acts as a tonic inhibitor of adrenocortical function, and its effects appear to be restricted to the cells in the zona glomerulosa, where it acts specifically to inhibit the secretion of the principal mineralocorticoid, aldosterone.

In Vivo Studies

In 1975, it was observed that the DA_2 receptor agonist bromocriptine inhibited the plasma aldosterone response to furosemide (a diuretic) in humans. This was the first evidence to indicate a role for dopamine as an inhibitor of aldosterone secretion *in vivo*. Further studies in humans then demonstrated that administration of metoclopramide, a nonselective DA receptor antagonist, led to an increase in plasma aldosterone concentration, which was independent of changes in PRA and circulating sodium, potassium, and ACTH concentrations. The stimulatory effect of metoclopramide on aldosterone secretion in humans can be inhibited by the orally active DA agonist ibopamine, which by itself acts as an inhibitor of aldosterone secretion. In addition, experimental evidence suggested that alterations in angiotensin-induced aldosterone responsiveness to changes in dietary sodium chloride may be related to dopaminergic modulation. Although these studies indicate that dopamine can act as a tonic inhibitor of aldosterone secretion, it is not clear whether these effects are mediated primarily by the central nervous system (because the DA agonists and antagonists used can cross the blood–brain barrier) or whether they may relate to more direct actions on the cells in the adrenal zona glomerulosa.

In Vitro Studies

A direct inhibitory action of dopamine on angiotensin II-induced aldosterone secretion was first reported in 1979 by McKenna, who used isolated bovine zona glomerulosa cells. Pharmacological experiments using rat adrenal zona glomerulosa homogenates later gave evidence for the presence of DA_1 receptors, which were coupled to activation of the G_s-protein, leading to the stimulation of cyclic AMP, and DA_2 receptors, which appeared to couple to the G_i-protein, causing a decrease in cyclic AMP. Compelling evidence suggests that the inhibitory effect of dopamine on aldosterone secretion is mediated by DA_2 receptors, which, when activated, inhibit intracellular Ca^{2+} concentrations or reduce transmembrane Ca^{2+} fluxes within zona glomerulosa cells. In particular, dopamine reduces angiotensin II-stimulated Ca^{2+} influx and angiotensin II-induced inositol phosphate production in rat zona glomerulosa cells. In addition, dopamine inhibits voltage-dependent T-type channels in cultured rat zona glomerulosa cells. No molecular studies have been reported yet on DA_2 receptors in the adrenal cortex, although DA_1 receptors have been shown to be present using *in situ* hybridization techniques to localize DA_1 receptor mRNA and immunohistochemistry to localize DA_1 receptor protein. The function of DA_1 receptors in the adrenal cortex requires further clarification, although tissue culture experiments in the rat zona glomerulosa suggest that activation of this receptor could lead to a stimulatory effect on aldosterone secretion.

Adrenocortical Sources of Dopamine

The concentrations of dopamine that are required to inhibit aldosterone secretion are far in excess of the concentrations of dopamine that circulate in plasma. Although dopamine is synthesized in the adrenal medulla, there is still no clear evidence that dopamine secreted by the adrenal medulla can affect steroid secretion in the zona glomerulosa directly. However, the enzyme L-AAAD, which catalyzes the conversion of circulating L-dopa to dopamine, has been located in the adrenal zona glomerulosa, and there is good evidence to suggest that circulating L-dopa may be the source of adrenocortical dopamine. The effects of dopamine within the adrenal cortex may therefore relate to a paracrine control mechanism, which is mediated by the activity of L-AAAD and the activation of specific DA receptors within the adrenal zona glomerulosa.

Conclusions

The principal actions of dopamine in the kidney are to increase renal blood flow and to promote natriuresis. The natriuretic effects of dopamine may be reinforced by its inhibitory action on aldosterone secretion in the zona glomerulosa cells of the adrenal cortex. These effects of dopamine are, to a great extent, mediated by the enzyme L-AAAD, which is located in the renal cortex and in the adrenal cortex.

The factors that control the expression and function of L-AAAD within the kidney and adrenal cortex, and the further development of selective DA_1 and DA_2 receptor agonists and antagonists, will be of paramount importance to our future understanding of the dopaminergic control of renal and adrenocortical function.

See Also the Following Articles

Adrenal Cortex; Aldosterone and Mineralocorticoid Receptors; Dopamine, Central.

Further Reading

Aherne, A. M., Vaughan, C. J., Carey, R. M. and O'Connell, D. P. (1997). Localization of dopamine D_{1A} receptor protein and messenger ribonucleic acid in rat adrenal cortex. *Endocrinology* **138**, 1282–1288.

Buu, N. T. and Lussier, C. (1990). Origin of dopamine in the rat adrenal cortex. *American Journal of Physiology* **258**, F287.

Goldberg, L. I. (1972). Cardiovascular and renal actions of dopamine: potential clinical applications. *Pharmacology Reviews* **24**, 1–29.

Goldberg, L. I., Volkman, P. H. and Kohli, J. D. (1978). A comparison of the vascular dopamine receptor with other dopamine receptors. *Annual Review of Pharmacology and Toxicology* **18**, 57–79.

Kebabian, J. W. and Calne, D. B. (1979). Multiple receptors for dopamine. *Nature (London)* **277**, 93–96.

Lee, M. R. (1982). Dopamine and the kidney. *Clinical Science* **62**, 439–448.

Lee, M. R. (1993). Dopamine and the kidney: ten years on. *Clinical Science* **84**, 357–375.

Missale, C., Lombardi, C., DeCotiis, R., et al. (1989). Dopaminergic receptor mechanisms modulating the renin-angiotensin system and aldosterone secretion: an overview. *Journal of Cardiovascular Pharmacology* **14**, S29.

Muller, J. (1987). *Regulation of aldosterone biosynthesis, physiological and clinical aspects*. New York: Springer Verlag.

Williams, B. C. (1986). Dopamine in the kidney. In: Winlow, W. & Markstein, R. (eds.) *The neurobiology of dopamine systems*, pp. 385–401. Manchester, UK: Manchester University Press.

Reproduction, Effects of Social Stress On

C A Shively
Wake Forest University School of Medicine, Winston-Salem, NC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C A Shively, volume 3, pp 360–365, © 2000, Elsevier Inc.

Social Stress and the Likelihood of Conception
Social Stress and Pregnancy Outcome
Long-Term Health of the Offspring

Glossary

Functional hypothalamic chronic anovulation	Inhibition of ovulation due to stress.
Hypothalamic-pituitary-adrenocortical axis	A major stress reactive system in the body involving the hypothalamus, pituitary, and adrenal glands.
Hypothalamic-pituitary-gonadal axis	The pathway of the reproductive system that controls gonadal function.
Luteal phase deficiency	Abnormalities of the luteal phase of the menstrual cycle that may include reduced sex hormone concentrations and changes in phase length.
Stereotypy	A locomotor (e.g., pacing) or self-directed (e.g., biting) behavior that is performed precisely the same way, repeatedly in quick succession.

Reproductive failure rates are uniformly high across the mammalian order. From the perspective of human medicine, reproductive failure is often viewed as evidence of a disorder that requires treatment. However, mammologists and evolutionary biologists understand reproductive failure as an evolutionary safeguard against wasting energy on reproduction when the likelihood of a successful outcome is relatively low. From this perspective, it makes sense that the reproductive system would be especially sensitive to environmental variables that might influence the outcome of reproductive effort. Female mammals have a greater investment in reproduction than males, and the female reproductive system is particularly sensitive to environmental factors signaling adverse conditions for successful reproduction. Thus, there is less data available

concerning stress effects on reproductive function of males than on that of females. For species that live in groups, social factors are key aspects of the environment. One type of social environmental influence to which the reproductive system of mammals appears quite sensitive is social stress. Social stress is defined here as a social environmental factor that perturbs homeostasis and elicits behavioral or physiological change. The purpose of this article is to examine the evidence of parental social stress on reproductive health, including fertility, pregnancy outcome, neonatal health, and the lifelong health of the offspring.

Social Stress and the Likelihood of Conception

Physiological Mechanisms

The mechanisms by which social stress may influence the likelihood of conception are not completely understood. However, stress-induced activation of the hypothalamic-pituitary-adrenocortical (HPA) axis inhibits the hypothalamic-pituitary-gonadal (HPG) axis at multiple levels. Corticotropin-releasing hormone (CRH) inhibits hypothalamic gonadotropin-releasing hormone (GnRH). CRH stimulates pituitary adrenocorticotropin (ACTH) secretion, which in turn increases adrenal glucocorticoid secretion. Glucocorticoids inhibit GnRH secretion, as well as pituitary luteinizing hormone and ovarian estrogen and progesterone secretion. CRH also stimulates β-endorphin, which in turn inhibits GnRH secretion. Gonadal dysfunction results from alterations in GnRH pulsatility from the hypothalamic pulse generator impairing ovulation. Other neurotransmitters that mediate stress effects on the GnRH pulse generator include dopamine, neuropeptide Y, norepinephrine, γ-aminobutyric acid (GABA), and endogenous opioid peptides.

Dopamine receptors appear to mediate stress-induced transcription of hypothalamic CRH and proopiomelanocortin, from which a number of peptides are derived, including ACTH and endogenous opioids. The results of a series of nonhuman primate studies suggest that dopamine is involved in the effect of one type of social stress on the reproductive system. A major social organizing mechanism of primate society is the social status hierarchy. Among group-living primates, low social status, or social subordination, appears stressful and is associated with lower reproductive output in females. Socially subordinate female cynomolgus monkeys (*Macaca fascicularis*) receive more aggression, maintain a high level of vigilance, are groomed less by conspecifics, and spend more time alone than their dominant counterparts. Their adrenal glands hypersecrete cortisol, and they are insensitive to glucocorticoid negative feedback. The central dopaminergic function of these females is also deranged, as evidenced by insensitivity to haloperidol, lower cerebrospinal fluid levels of homovanillic acid (HVA), a dopamine metabolite, and lower D_2 receptor binding potential in striatum observed by positron emission tomography. In males, the relationship between social subordination stress and reproductive output is less consistent than in females. Associations between social status and offspring sired are observed in some years but not others. Likewise, the relationship between low social status and activated HPA function is less apparent in male than in female Old World primates. However, central dopaminergic function of subordinate males is different than that of dominants, as evidenced by lower cerebral spinal fluid (CSF) HVA concentrations and lower striatal D_2 receptor binding potential. Thus, social subordination stress may alter central dopaminergic function, increase HPA activity, and inhibit HPG function.

Another mechanism that may mediate stress effects on conception involves reactive oxygen species (ROS). ROS can modify cell function and endanger cell survival. Because these are normal consequences of oxygen metabolism, ROS must be continuously deactivated by antioxidant defense mechanisms to avoid oxidative stress. There is some evidence that psychosocial stress may increase ROS, decrease DNA repair, and alter rates of apoptosis. Thus, oxidative damage may occur during stress. High levels of ROS are found in the semen of infertile men and the peritoneal and follicular fluid of infertile women.

Inhibition/Delay of Sexual Maturation

Two hypotheses address the relationship between social stress and sexual maturation. The social suppression hypothesis posits that stress suppresses sexual development. In support of this hypothesis, it has been widely observed in monkeys that low social status is associated with delayed sexual maturation and that dominant females reproduce at an earlier age than subordinate females. Likewise, girls from higher social classes enter puberty earlier than girls from lower social classes. Delayed sexual maturation following severe stress (e.g., war) has been documented in boys and girls, though is not well studied. In contrast, the psychosocial acceleration theory posits that under certain types of stressful conditions in which low-quality parental investment signals that parental investment does not promote reproductive success, sexual maturation may accelerate so that reproduction may occur earlier and more often. There is some evidence that family warmth and positivity may

delay, and family conflict and coercion may accelerate, pubertal development. Thus, the nature of the psychosocial stressor may be critical in predicting effects on sexual development.

Inhibition/Delay of Ovulation

Low social status is associated with inhibition of ovulation in baboons and macaques (*Macaca* spp.). Female cynomolgus monkeys (*M. fascicularis*) have menstrual cycles that are very similar to those of women. Socially subordinate female cynomolgus monkeys have low luteal phase progesterone concentrations, low follicular phase estradiol concentrations, a greater proportion of menstrual cycles with impaired luteal function, and a greater proportion of anovulatory menstrual cycles. Furthermore, when female monkeys change social status, their menstrual cycle characteristics change also (Table 1), suggesting that the reproductive system of these females is exquisitely sensitive to their social environment.

Socially suppressed ovarian function in these monkeys is associated with deleterious effects on the skeletal and cardiovascular systems, suggesting that socially induced hormonal impairments unaccompanied by amenorrhea have far-reaching effects on health.

Social stress may also inhibit ovulation in women, a condition referred to as functional hypothalamic amenorrhea or functional hypothalamic chronic anovulation (FHCA). Diagnosis is by exclusion of all other potential organic causes. The best indicator that FHCA is stress related is that it is generally accompanied by hypercortisolemia. Note that socially subordinate female monkeys that are stressed do not necessarily have more anovulatory cycles than their nonstressed dominant counterparts. Thus, in extrapolating to women, it is important to consider that stress-induced suppression of reproductive function may not manifest as frank amenorrhea, complicating diagnosis. In women, perturbations in GnRH pulsatility may result in ovarian dysfunction manifested as amenorrhea, polymenorrhea, oligomenorrhea, or luteal phase deficiency, the latter of which may go undetected unless infertility issues arise. FHCA also appears reversible in women. The prevalence of FHCA varies depending on the subpopulation considered. Estimates in the general population range from 2 to 7%, whereas in high-risk populations, such as women going to college or joining the armed forces and women with depression, the prevalence may be as high as 70%. Under these circumstances ovulatory impairment appears to coincide with psychosocial stress and precede nutritional stress. Women that develop FHCA are more likely to be anxious, be perfectionists, and have difficulty coping.

Impairment of Hypothalamic-Pituitary-Testicular Function

During periods of social instability, low social status and aggression received are both associated with suppressed plasma testosterone concentrations in male macaques and baboons. In men, stressful events (e.g., work stress, exams, or competitions) may decrease testosterone levels, which can lead to decreased libido, diminished muscle mass, change in hair growth, and decreased spermatogenesis. Decreased testosterone may be due to perturbations in GnRH pulsatility or to direct effects of cortisol on testicular Leydig cell testosterone production. Social stress (e.g., death of a close family member) can result in oligoasthenozoospermia, which may remain undetected unless infertility issues arise. Decreased sperm number and quality are associated with decreased fertility.

Social Stress and Pregnancy Outcome

Physiological Mechanisms

There are several ways in which neuroendocrine stress responsivity may affect pregnancy. CRH regulates the HPA responses to stress. During pregnancy

Table 1 Effect of changing social status on menstrual cycles[a]

Initial status	*Subordinate*		*Dominant*		*Final status effect p ≤*
Final status	*Subordinate*	*Dominant*	*Subordinate*	*Dominant*	
Number of cycles	15.2 (0.94)	15.4 (1.03)	14.4 (1.52)	15.1 (1.1)	0.60
Percent ovulatory	68.1 (9.19)	91.4 (3.36)	81.6 (6.34)	88.4 (4.42)	0.03
Percent impaired	19.4 (5.21)	4.2 (1.65)	8.4 (2.8)	8.1 (2.60)	0.08
Percent anovulatory	12.5 (6.00)	4.5 (2.21)	9.9 (5.44)	3.5 (2.76)	0.17
Progesterone (ng/ml)	6.7 (0.96)	10.4 (1.72)	8.1 (0.98)	9.6 (0.97)	0.03

[a] Means and standard errors. Social status of 42 female cynomolgus monkeys living in small social groups was documented for 2 months. The constituency of the social groups was changed such that half of the subordinates became dominant and half of the dominants became subordinate, and social status and menstrual cycle characteristics were documented for 24 months.

it also coordinates and controls the physiology of parturition. CRH is produced in the placenta, and placental CRH is stress sensitive. Maternal pituitary-adrenal hormones stimulate placental CRH secretion. Maternal stress is positively associated with circulating ACTH concentrations, and maternal social support is negatively associated with circulating concentrations of ACTH, β-endorphin, and cortisol. Maternal β-endorphin secretion during the third trimester is associated with increased fetal heart rate reactivity and decreased uroplacental blood flow, the latter being consistent with fetal hypoxia. Elevated CRH levels are associated with preterm labor and preterm delivery. Onset of the birthing process also involves oxytocin, vasopressin, and prostaglandins, as well as HPA activation. Secretion of these hormones may be altered by stress, thus affecting the timing of birth. Indeed, uteri from preterm deliveries appear to be hypersensitive to oxytocin. Stress-induced catecholamine and oxytocin may reduce the threshold of uterine contractility, thus resulting in preterm labor.

There is substantial evidence that HPA axis peptides (including ACTH and β-endorphin) affect the development of fetal brain and behavior and that perinatal exposure to ACTH and its analogues permanently alters growth and behavior. Thus, maternal stress may alter the trajectory of development beginning *in utero*.

Neuroendocrine stress responses are known to suppress the immune system; however, these processes have been little studied during pregnancy. Pregnancy involves immunosuppression in order to avoid fetal rejection. Thus, stress-related immunosuppression may be superimposed on the normal physiological immunosuppression of pregnancy, potentially affecting pregnancy outcome. Stress-induced immunosuppression increases vulnerability to infection. Infection, in turn, may induce preterm delivery. Preeclampsia is also associated with altered immune function and thus may be affected by stress-induced immunosuppression. Preterm birth is associated with enhanced expression of pro-inflammatory cytokines. Pro-inflammatory cytokines stimulate prostaglandin and metalloprotease synthesis in gestational tissues, which promote spontaneous labor and rupture of membranes. They also stimulate fetal-placental production of steroids and CRH, which promote parturition.

Norepinephrine and epinephrine are stress-responsive catecholamines that decrease blood flow to the uterus. Decreased uterine blood flow may cause hypoxia, hypotension, bradycardia, and malformations in the fetus. Hypoxia may result in increased ROS production. Oxidative stress due to increased ROS and/or decreased antioxidant defense mechanisms has been implicated in miscarriage and preeclampsia.

Thus, maternal, placental, and fetal neuroendocrine, immune, and cardiovascular processes link maternal stress during pregnancy with adverse pregnancy outcomes.

Failure of Implantation

It is difficult to determine failure of implantation, especially to discriminate it from ovulatory difficulties. Both social status and current ecological conditions affect the likelihood of successful implantation in yellow baboons. Under poor environmental conditions, socially dominant females (with low social stress) are more likely to achieve successful implantation than their subordinate, stressed counterparts. Likewise, in an *in vitro* fertilization study, implantation rates were reportedly higher in women with relatively low cardiovascular reactivity to a standardized laboratory stressor.

Spontaneous Abortion

There is an association between spontaneous abortion and psychological stress in pigtail macaques (*M. nemestrina*). Low social status, psychological stress, low control, and low social support have been associated with spontaneous abortion in humans. In 192 women with recent spontaneous abortions, it was found that recent negative life events significantly increased the likelihood of spontaneous abortion of chromosomally normal conceptions. In perhaps one of the largest studies available ($n = 214{,}108$), adverse pregnancy outcomes, including spontaneous abortion, were most frequent in women working in jobs characterized by high demand and low control. In another study of several thousand women, stressful work increased spontaneous abortion by a factor of 2.45 (95% confidence interval [CI] 1.03–5.81) in older women, 2.96 (95% CI 1.16–7.52) in smokers, and 2.27 (95% CI 0.97–5.27) in primigravid women. Thus, stressful life events, work characterized by high demand and low control, and stressful work may increase rates of spontaneous abortion.

Maternal Stress and Neonatal Mortality

Associations of neonatal mortality and maternal social stress are well documented in nonhuman primates. Rhesus macaque infants born to socially dominant females have a higher rate of survival than those born to socially subordinate mothers. Higher mortality also has been observed among offspring of psychologically stressed pigtail macaques. Chimpanzee and gorilla offspring born to subordinate mothers are subject to greater neonatal mortality than those born to dominants, in part due to infanticide. Maternal psychological stress and low social status have

been associated with neonatal mortality in humans. In a large study ($n = 214{,}108$), adverse pregnancy outcomes, including stillbirth, were most frequent in women working in jobs characterized by high demand and low control. Maternal risk factors for sudden infant death syndrome (SIDS) include low socioeconomic status (SES), unmarried status, teen age, less education, depression, drug use, and ethnicity. These characteristics are also risk factors for low birth weight, and low birth weight is a powerful predictor of SIDS. While it is difficult at this time to tease apart the relative contributions of these risk factors, all may involve increased maternal stress during gestation.

Maternal Stress and Neonatal Morbidity

The effects of maternal social stress during pregnancy on neonatal morbidity are more subtle and difficult to determine than neonatal mortality. Measures of neonatal morbidity include labor and delivery complications, gestational age, Apgar scores, birth weight, and neuromotor and neurobehavioral development during the first 30 days.

A series of studies in squirrel (*Saimiri boliviensis*) and rhesus monkeys demonstrated that mild maternal stress (e.g., removal from a home cage daily followed by 10 min of white noise during mid-late gestation) results in lower birth weight, impaired neuromotor development (as measured by a modified Brazelton Newborn Behavioral Assessment Scale), and impaired neurobehavioral development of attention and orienting. Early gestation stress was associated with more profound impairments to growth and neuromotor development during the first postnatal month than late gestation stress. These impairments are mimicked by administering ACTH to pregnant females for 2 weeks at mid-pregnancy, suggesting that activation of the HPA axis is instrumental in the deleterious effects of maternal prenatal stress on infant morbidity. Similar decrements in offspring development continue into adolescence (see later).

A large number of studies have been conducted to test the hypothesis that prenatal maternal stress causes infant morbidity in humans, and the balance of the literature supports this hypothesis. Birth weight and gestational age are the most commonly measured outcome variables. Low birth weight can be due to preterm delivery (less than 37 weeks gestation with an appropriate infant weight for gestational age) or to retarded fetal growth (inappropriate weight for gestational age). Maternal anxiety and depression result in low birth weight and perhaps smaller head size. The magnitude of the effect is comparable to that of smoking. Composite measures of maternal prenatal stress (including some combination of trait or state anxiety, depression, life events, social support, and self esteem) best predict preterm delivery/low gestational age. Likewise, the association between psychosocial stressors and birth weight is strongest when composite measures of maternal stress (e.g., exposure to stressors such as life event scales, social support, anxiety, health risk behaviors) were used. The timing of the stress is important. Low birth weight and shorter gestational age have been observed when the stressors occur in the first trimester, whereas preterm delivery is associated with third trimester stress. In one of the largest studies available ($n = 214{,}108$), adverse pregnancy outcomes, including light-for-date birth weight and full-term low birth weight, were most frequent in women working in jobs characterized by high demand and low control.

Long-Term Health of the Offspring

Maternal Stress and Offspring Central Nervous System (CNS) Function

The effects of maternal stress during gestation have been shown to affect CNS function of rhesus monkeys (*M. mulatta*) through adolescence (4 years of age). Six-month-old offspring from control mothers and mothers that experienced a daily mild stressor (three random white noise bursts over 10 min) were evaluated during the latter half of pregnancy. In a response to a novel environment, offspring of prenatally stressed mothers exhibited more disturbance behavior (e.g., clinging to attachment object, self-directed stereotypic behaviors) and lower levels of gross motor and exploratory behavior than those of nonstressed mothers. At 8 and 18 months, prenatal stress-derived monkeys had higher CSF concentrations of metabolites of dopamine and norepinephrine, suggesting altered CNS function. The behavioral responses of these monkeys to social separation stress were evaluated. Disturbances in exploratory, locomotor, and clinging behavior observed at 6 months of age continued, and disturbances in social play were also observed. The effects of prenatal stress on later behavioral responses to stress were more profound for social than for nonsocial behavior, and suggest that offspring from prenatally stressed mothers are behaviorally hyperresponsive to stress. Between 6 months and 1 year of age, male offspring of prenatally stressed mothers had smaller, and females had larger, corpus callosum. At 14–19

months of age, prenatal stress-derived monkeys had higher ACTH and cortisol levels at baseline and under anesthesia. Furthermore, in response to stress, prenatal stress-derived offspring had higher ACTH and similar cortisol responses as control offspring. These observations suggest that the HPA axis of offspring of primate mothers stressed during pregnancy hyperresponds to stress later in life, and that the components of this stress-responsive system are dysregulated. The behavior of these prenatal stress-derived offspring was characterized again at 4 years of age, a time corresponding to human adolescence. At this age, the behavioral pathology of the prenatal stress-derived offspring was profound and included greater locomotion, stereotypies (pathological repetitive behaviors such as rocking), clinging, and self-clasping. These monkeys explored less in a novel environment and engaged in less social interaction in a playroom environment. At 5–7 years of age, *in vivo* imaging studies revealed a perturbed striatal dopaminergic system, which is critically involved in drug abuse vulnerability and sensitive to adult social stresses. Taken together, these observations suggest that prenatal stress can have profound and long-term effects on the behavior and neurobiology of primate offspring that is particularly apparent in response to stress.

Similar data from human studies are difficult to acquire. However, the results of a number of studies suggest that prenatal stress may perturb CNS function of newborns. There are reports of increased fetal activity associated with maternal anxiety and catastrophic stress due to earthquake. Anxious, type A, or depressed mothers have newborns who cry more and are more difficult to soothe. First trimester maternal stress and anxiety have been associated with low scores on cognitive and psychomotor tests, poor adaptation to novel environments, and difficult behavior in infants 8 months of age. At 9 years of age, attention deficits and aggressiveness were more common in the offspring of mothers who were anxious during pregnancy. Maternal stress is also linked to offspring psychopathology. Schizophrenia, in particular, has been linked to prenatal stress. Maternal exposure to death of a loved one, influenza infection, or catastrophic stressors increase the risk of the offspring developing schizophrenia, the onset of which is most common between 16 and 30 years of age.

The Significance of Low Birth Weight to Health in Later Life

Cardiovascular disease (CVD) is the leading killer ofmen and women. Low birth weight has been associated with all-cause mortality, coronary heart disease, stroke, type II (non-insulin-dependent) diabetes mellitus, hypertension, and the metabolic syndrome. The metabolic syndrome is a cluster of characteristics (e.g., abdominal fat deposition, hyperglycemia, hyperinsulinemia, hypertension, and an atherogenic lipid profile) that greatly increase risk of CVD. Given the patterning of associations, it appears that low birth weight may be associated with lifelong derangement of the metabolic function of several major systems that culminates in increased risk of cardiovascular disease. Hypotheses currently being tested to explain fetal programming of adult health include the possibility that suboptimal environmental influences on fetal growth lead to permanent changes in organogenesis (i.e., particularly those involved with energy metabolism) that may serve to preserve the growth of certain organs, like the brain, to the expense of the function of other systems. Prenatal exposure to excess cortisol is considered a likely mechanism by which maternal stress during gestation may lead to fetal programming of cardiovascular, metabolic, and neuroendocrine disorders in adult life.

Exaggerated maternal stress responses can bathe the fetus in excessive glucocorticoids, resulting in an adult programmed for exaggerated HPA responses to stress. If this adult gestates young, she may also produce offspring with exaggerated stress responses, and thus intergenerational transmission of stress reactivity and subsequent ill health may be perpetuated. Indeed, low-birth-weight women are more likely to have low-birth-weight babies.

See Also the Following Articles

Childbirth and Stress; Chronic Social Stress: GR Sensitivity in Leukocytes; Social Stress, Animal Models of; Stress Induced Anovulation.

Further Reading

Amiel-Tison, C., Cabrol, D., Denver, R., et al. (2004). Fetal adaptation to stress: Part II. Evolutionary aspects; stress-induced hippocampal damage; long-term effects on behavior; consequences on adult health. *Early Human Development* **78**(2), 81–94.

Berga, S. L. (1997). Behaviorally induced reproductive compromise in women and men. *Seminars in Reproductive Endocrinology* **15**, 47–53.

Buitelaar, J. K., Huizink, A. C., Mulder, E. J., et al. (2003). Prenatal stress and cognitive development and temperament in infants. *Neurobiology of Aging* **24**(supplement 1), S53–S60.

Burton, G. J. and Jauniaux, E. (2004). Placental oxidative stress: from miscarriage to preeclampsia. *Journal of the Society for Gynecologic Investigation* **11**(6), 342–352.

Drake, A. J. and Walker, B. R. (2004). The intergenerational effects of fetal programming: non-genomic mechanisms for the inheritance of low birth weight and cardiovascular risk. *The Journal of Endocrinology* **180**(1), 1–16.

Federenko, I. S. and Wadhwa, P. D. (2004). Women's mental health during pregnancy influences fetal and infant developmental and health outcomes. *CNS Spectrum* **9**, 198–206.

Huizink, A. C., Mulder, E. J. and Buitelaar, J. K. (2004). Prenatal stress and risk for psychopathology: specific effects or induction of general susceptibility? *Psychological Bulletin* **130**, 115–142.

Kaplan, J. R. and Manuck, S. B. (2004). Ovarian dysfunction, stress, and disease: a primate continuum. *Institute of Laboratory Animal Resources Journal* **45**, 89–115.

Koenig, J. I., Kirkpatrick, B. and Lee, P. (2002). Glucocorticoid hormones and early brain development in schizophrenia. *Neuropsychopharmacology* **27**(2), 309–318.

Mulder, E. J., Robles de Medina, P. G. and Huizink, A. C. (2002). Prenatal maternal stress: effects on pregnancy and the (unborn) child. *Early Human Development* **70**(1-2), 3–14.

Roberts, A. D., Moore, C. F., DeJesus, O. T., et al. (2004). Prenatal stress, moderate fetal alcohol, and dopamine system function in rhesus monkeys. *Neurotoxicology and Teratology* **26**(2), 169–178.

Schneider, M. L., Roughton, E. C., Koehler, A. J., et al. (1999). Growth and development following prenatal stress exposure in primates: an examination of ontogenetic vulnerability. *Child Development* **70**(2), 263–274.

Seckl, J. R. (2004). Prenatal glucocorticoids and long-term programming. *European Journal of Endocrinology* **151**(supplement 3), U49–U62.

Sullivan, F. M. and Barlow, S. M. (2001). Review of risk factors for sudden infant death syndrome. *Paediatric and Perinatal Epidemiology* **15**(2), 144–200.

Wasser, S. K. and Barash, D. P. (1983). Reproductive suppression among female mammals: implications for biomedicine and sexual selection theory. *The Quarterly Review of Biology* **58**, 513–538.

Zitzmann, M. and Nieschlag, E. (2001). Testosterone levels in healthy men and the relation to behavioural and physical characteristics: facts and constructs. *European Journal of Endocrinology* **144**(3), 183–197.

Reproductive Dysfunction in Primates, Behaviorally Induced

J L Cameron
University of Pittsburgh and the Oregon Regional Primate Research Center, Beaverton, OR, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 366–372, © 2000, Elsevier Inc.

Overview
Stress-Induced Impairment of the Hypothalamic-Pituitary-Testicular Axis in Male Primates
Stress-Induced Impairment of the Hypothalamic-Pituitary-Ovarian Axis in Female Primates
Social Status as a Determinant of Reproductive Capacity
Mechanisms Underlying Stress-Induced Reproductive Dysfunction in Primate Species
Individual Susceptibility to Stress-Induced Suppression of Reproductive Function
Summary

Glossary

Hypothalamic-pituitary-adrenal axis	The complex of (1) specialized cells in the hypothalamus that secrete the neuroendocrine peptide, corticotropin-releasing hormone, (2) cells in the anterior pituitary that secrete adrenocorticotropin (ACTH), and (3) the adrenal gland that secretes glucocorticoid and mineralocorticoid hormones in response to ACTH.
Endogenous opioid peptides	Peptides released both within the brain and from pituitary cells that inhibit binding of morphinelike compounds to receptors.
Individual variability in stress responsiveness	Difference in physiological responses among individuals in a particular species to a given stress.

Reproductive hormones	Pituitary gonadotropins, including luteinizing hormone and follicle-stimulating hormone, and gonadal steroid hormones, including estradiol, progesterone, and testosterone.
Social stress	Stress arising from interactions with others or the perception of other's attitudes.
Social status	Placement of a particular animal within a social hierarchy, with subordinate animals deferring to the wishes of more dominant animals.

Both acute and chronic psychological and social stresses can impair reproductive hormone secretion in a variety of nonhuman primate species. This impairment can be subtle, consisting of a mild suppression in reproductive hormone secretion, or can be dramatic, causing a complete suppression of fertility and reproductive behavior. Among individuals there are marked differences in the responsiveness of the reproductive axis to social stresses, with factors contributing to this variability including the type of stress experienced, the magnitude and duration of stress, the perception of the stress by the individual, the social status of the individual, the concurrent level of aggressive behavior displayed by the individual, seasonal cues, and the prior level of activity within the reproductive axis (Figure 1). Mechanisms contributing to social stress induction of reproductive dysfunction include activation of the adrenal axis, increased secretion of endogenous opioids, increased prolactin release, and changes in sensitivity to gonadal steroid hormone feedback; however, in most forms of social stress the pathways leading to reproductive dysfunction remain to be elucidated.

Overview

While many forms of physical stress, such as energy restriction, temperature stress, infection, pain, and injury, have been clearly associated with the impairment of reproductive function, the effects of behaviorally induced stresses, i.e., psychological and social stresses, on the activity of the reproductive axis have been less well studied. The topic of behaviorally mediated stresses on reproductive function is of particular interest in primate species, which live in complex social groups and have higher cortical brain areas similar to humans, making study of these species particularly useful in understanding how such stresses may impact on reproduction and fertility in humans. This article provides a review of what is known currently about the role of psychological and social stresses in modulating the activity of the reproductive axis in nonhuman primate species. The review will show that the impact of psychological and social stresses on reproductive function is associated with significant variability among individuals and will detail the factors known to influence this variability in stress responsiveness.

Stress-Induced Impairment of the Hypothalamic-Pituitary-Testicular Axis in Male Primates

A variety of acute psychological stresses have been shown to cause the acute suppression of circulating luteinizing hormone (LH) and testosterone levels in male nonhuman primates. Examples include studies showing that the restraint of male rhesus monkeys suppresses pulsatile LH and testosterone release, and that mean testosterone concentrations decline rapidly in male rhesus monkeys in response to aggressive attacks by other monkeys. However, in other studies

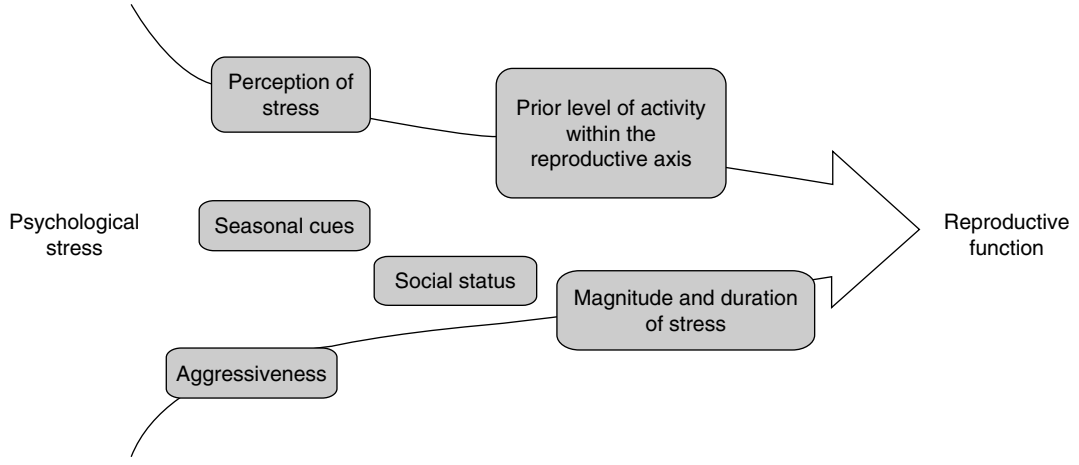


Figure 1 Schematic diagram of variables important in determining whether psychological stress will stimulate or inhibit reproductive function within an individual.

the acute response to stress in male primates has been shown to be more variable. In baboons, darting and capture have been shown to lead to an immediate suppression of testosterone in some animals, but to an initial increase in plasma testosterone levels followed by a much more subtle decline in other animals. Similarly, in response to pairing with an unfamiliar male animal, some male squirrel monkeys have an acute release of testosterone, whereas other males do not show this initial rise in testosterone. In this case, monkeys showing acute testosterone release are the behaviorally dominant monkeys in the pairs, whereas subordinate monkeys fail to show the acute testosterone release. Although many acute psychological stressors can suppress reproductive hormone secretion in male primates, not all such stresses have this effect. For example, our laboratory has used the acute psychological stress of exposing male rhesus monkeys to the sight of leather capture gloves to examine the physiological sequelas of acute psychological stress and has shown that although heart rate and cortisol release are elevated acutely in this paradigm and that immune cell function is suppressed for a number of hours, LH and testosterone secretion are not influenced. Although this stress evokes physiological stress responses in several systems, it seems likely that monkeys perceive this stress as being less threatening than a stress such as an aggressive attack by other monkeys. Thus, the perception of severity of stress is likely to be important in determining whether the stress will lead to a suppression of reproductive hormone secretion.

A number of chronic social stresses have also been associated with a marked and sustained suppression of reproductive hormone secretion in males, again with several investigations showing a marked dichotomy in response between dominant versus subordinate males, and other investigations showing that variability in the response of the reproductive axis to stress is dependent on environmental factors. Examples of prolonged stresses shown to lead to decreased testosterone secretion include social stress associated with loss or removal of the dominant male in the social group, housing males in all male social groups, and housing males in isolation. Interestingly, several social factors appear to modulate the response of the reproductive axis to stress in such conditions of chronic social stress. In a number of studies, a high degree of correlation between aggressiveness and testosterone titers, with the more aggressive males showing higher circulating levels of testosterone, has been reported. There also appears to be an interaction between the social status and the level of environmental stress, such that in periods of social stability there can be no difference in plasma testosterone levels between dominant and subordinate animals, but in periods of social instability the dominant animals can show higher plasma testosterone levels than the subordinate males.

Based on the ability of both acute and chronic social stress to suppress reproductive hormone secretion, it has been posited that social status (i.e., dominance rank) plays an important role in determining the lifetime reproductive success in primates, with subordinate animals experiencing a greater degree of social stress and having a lesser degree of reproductive success. However, evidence for this hypothesis is weak, and it would appear that again multiple factors, including current dominance rank, time of year, magnitude of stress, aggressiveness of the animal, and level of activity of the reproductive axis prior to stress exposure, can all play roles in modulating reproductive success, in the same way these factors influence reproductive hormone secretion. Thus there are a number of instances in which subordinate males have similar long-term reproductive success when compared to dominant males.

Stress-Induced Impairment of the Hypothalamic-Pituitary-Ovarian Axis in Female Primates

Chronic social stress can also impair reproductive function in female nonhuman primates, which often show an increase in the length of the menstrual cycle and an increased incidence of anovulation in response to being maintained in stressful situations. In general, stress-induced lengthening of the cycle is characterized predominantly by a lengthened follicular phase and a normal length luteal phase. Social stresses associated with the inhibition of reproductive function include moving females to new social groups, housing in isolation, and even moving singly housed females to a new room where they can only interact with other animals through visual, auditory, and olfactory cues.

A number of studies have shown that dominant females have greater reproductive success than subordinate females in various nonhuman primate species. It has been widely assumed that subordinate females experience more stress than dominant females and that this stress plays a role in the relative suppression of reproductive function. This assumption has been born out by studies showing that subordinate female macaques housed in small group settings have larger adrenals and an enhanced adrenal responsiveness to adrenocorticotropin (ACTH) stimulation accompanying an increased incidence of anovulatory cycles. However, as in males, evidence shows that dominance rank may only determine reproductive success

when environmental resources are limiting, at least in some species. This supports the notion that a number of environmental and social factors play a role in determining the impact of stress exposure on reproductive function.

Social Status as a Determinant of Reproductive Capacity

In the preceding paragraphs, evidence has been presented showing that exposure to stress can suppress reproductive hormone secretion in nonhuman primate species that show normal adult reproductive function in nonstressed conditions. It seems likely that the ability of stress to impair reproductive function in these animals may well be analogous to the ability of psychological stress to impair reproductive function in humans. In addition, however, there are several species of nonhuman primates in which social interactions play a more absolute role in determining the activity of the reproductive axis, and this appears to be particularly true in the case of females. Specifically, in talpoin monkeys (*Miopithecus talapoin*) and marmosets, only the dominant female in a social group is reproductively active and all subordinate females are reproductively inactive until removal of the dominant female from the social group. As discussed later, evidence shows that social stress cues play an important role in this form of regulation of the reproductive axis.

In social groups of common marmosets (*Callithrix jacchus*), only the dominant female is reproductively active and bears offspring. As the female offspring mature in their natal group they remain reproductively quiescent, essentially held in a prepubertal state. These subordinate females display no signs of ovulation, they have low circulating concentrations of LH and ovarian steroid hormones, and the ovaries remain dramatically smaller than those of the dominant female, containing only small follicles and stromal tissue. The infertility can be reversed by several weeks of treatment with a pulsatile gonadotropin releasing hormone (GnRH) regimen, indicating that the suppression of reproductive function is at a central level to block the GnRH drive to the reproductive axis. Similar findings have been reported in female talpoin monkeys.

In these species, subordinate males also appear to have a suppression of reproductive axis activity in comparison to the dominant males, but the impairment of the reproductive axis is not as severe in males as in females. In male marmosets, the subordinate animals living in social group settings have lower circulating concentrations of LH and testosterone. In male talpoin monkeys, there is a general trend for the most dominant males to display the greatest amount of sexual behavior and the most subordinate males to show little or no sexual behavior when housed in social groups. When housed in social groups of multiple males and females there is also a trend for the dominant males to have higher circulating concentrations of testosterone than the most subordinate males, but on a day-to-day basis this relationship between social rank and plasma testosterone concentrations is not always apparent. Interestingly, when male talpoin monkeys are moved to individual cages the correlation between social rank and plasma testosterone concentrations completely disappears. It thus appears that social rank modulates the activity of the reproductive axis, but that this effect is only apparent when monkeys are in a social setting where dominance hierarchies are in place.

The two most common routes by which social cues can suppress reproductive function in mammals are by behavioral intimidation and by pheromonal cues. Although there is some evidence that pheromonal cues play a role in the reproductive quiescence in the subordinate animals of these species, ablation of the olfactory epithelium and vomeronasal organ does not prevent suppression of reproductive function in subordinate animals, clearly indicating that other cues play a role in maintaining reproductive suppression. Behavioral intimidation of subordinate animals is apparent in both of these species. When marmosets are first placed in social groups there is a brief period where they receive an increased amount of harassment and intimidation from dominant animals, and in talpoin monkeys the place within the social hierarchy of a particular animal is established by the direction and frequency of aggressive interactions between various animals. The most dominant animals receive the least amount of aggression from others in the social group, whereas the most subordinate animals receive the most aggression. In addition, subordinate animals visually monitor more dominant animals at very frequent intervals. Subordinate animals also show an increase in circulating concentrations of the stress hormones, prolactin and cortisol. Cumulatively, data discussed earlier indicate that the stress associated with subordinate status within a social group at least partially underlies the suppression of reproductive function in these species of primates.

Mechanisms Underlying Stress-Induced Reproductive Dysfunction in Primate Species

Activation of the Adrenal Axis

In many cases, social stress-induced reproductive dysfunction is accompanied by an activation of the

hypothalamic-pituitary-adrenal axis, and it is possible that in these cases hormones of the adrenal axis may play a role in suppressing reproductive hormone secretion, and ultimately sexual behavior. Several secretory products of the adrenal axis have been shown to be capable of suppressing activity of the reproductive axis. Corticosteroids have been shown to both directly inhibit GnRH secretion and suppress the pituitary responsiveness to GnRH, although these effects of corticosteroids require high levels sustained over a several-week time course. Corticotropin-releasing hormone (CRH), the central hypothalamic-releasing factor of the adrenal axis, has also been shown to be a potent inhibitor of LH secretion when it is administered in pharmacological experiments. In addition, for some forms of physical stress, the administration of CRH antagonists can prevent the stress-induced suppression of reproductive hormone secretion.

Despite the evidence that elevated activity of the adrenal axis often accompanies suppressed reproductive hormone secretion in various forms of psychological and social stress in primates, there is no evidence to date that in these psychological stress conditions that activation of the adrenal axis is indeed causing the suppression of the reproductive axis. A causal relationship would only be established if it could be shown that blocking activation of the adrenal axis reversed the psychological stress-induced inhibition of the reproductive axis, and such experiments have not been performed to date. In addition, it is important to recognize that in at least several situations where other physical stresses are associated with both an increase in adrenal hormone secretion and a decrease in reproductive hormone secretion that such studies have shown that activation of the adrenal axis is not causing the suppression of the reproductive axis. Thus, the role that activation of the adrenal axis actually plays in causing suppression of the reproductive axis in any conditions of social or psychological stress remains to be elucidated.

Other Mechanisms

Evidence shows that the increased release of endogenous opioid peptides plays at least a partial role in suppressing the activity of the reproductive axis in some forms of social stress. There are several reports that administration of the opioid receptor antagonist, naloxone, can restore or partially restore LH secretion in conditions of psychological stress, including the suppression of LH and testosterone secretion in chair-restrained rhesus monkeys and after anesthetic darting in olive baboons. Naloxone can also restore LH secretion when administered to ovariectomized, subordinate marmosets. However, because opioid receptor antagonists are not effective in restoring activity of the reproductive axis in all forms of social stress-induced reproductive dysfunction, clearly other mechanisms are also involved in transmitting information about social stress to the reproductive axis.

Subordinate talpoin monkeys show a marked elevation in circulating prolactin levels compared to dominant animals, and it is possible that this increase in prolactin may play a role in suppressing the activity of the reproductive axis in this species, in that bromocryptine, a dopamine agonist, which reduces prolactin secretion, is able to at least partially restore reproductive hormone secretion. However, prolactin does not appear to mediate stress-induced suppression of the reproductive axis in all species. Even in another primate species in which subordinate animals show marked suppression of the reproductive axis, the marmoset, prolactin levels are not elevated and do not appear to play a role in mediating the inhibition of reproductive function.

In marmosets, evidence shows that the social suppression of reproductive function is associated with an increased sensitivity of the reproductive axis to estradiol negative feedback, as shown by the ability of low levels of estradiol to inhibit LH secretion in subordinate, but not dominant, female marmosets. In talpoin monkeys, removal of the subordinate female from the social group leads to a restoration of reproductive hormone secretion and increased sensitivity of the reproductive axis to estradiol-positive feedback. The neural mechanisms that underlie changes in sensitivity to estradiol feedback in these conditions remain unknown.

A review of all the studies examining social stress-induced suppression of the reproductive axis suggests that there are likely to be multiple neural systems that contribute to the inhibition of reproductive function in times of social and psychological stress. Clearly, there appear to be species differences in the neural pathways leading to reproductive inhibition. However, perhaps more interesting to consider is the possibility that within many species a similar set of pathways is involved in the suppression of reproductive axis activity, but that differential activation or inhibition of specific pathways within this complex is dependent on the specific stress experienced, as well as within a specific species. A great deal more work is needed in this field to dissect the neural mechanisms underlying behaviorally induced reproductive dysfunction and how this varies with the type of stress and with species.

Individual Susceptibility to Stress-Induced Suppression of Reproductive Function

A concept that arises again and again in this area of investigation is that in most all species some animals appear to be more susceptible to the stress-induced

suppression of reproductive function than others. As discussed in this article, subordinate social status is associated with a greater inhibition of the reproductive axis, either in response to an imposed stress or in the normal living situation, in a diversity of primate species, including rhesus monkeys, cynomolgus monkeys, baboons, talpoin monkeys, and marmosets. A difficult question to answer that needs to be addressed in future studies is whether these animals are intrinsically more susceptible to stress-induced reproductive dysfunction or whether they experience more stress and thus have greater inhibitory drive to their reproductive axis.

Summary

As detailed in this review, both short- and long-term exposure to a variety of psychological and social stresses can suppress reproductive hormone secretion and lead to a decrease in reproductive behaviors and fertility in a number of nonhuman primate species. Although there appear to be some species differences in the degree or route of inhibition of the reproductive axis caused by exposure to such stresses, there also appear to be several general principles that are common to most species. Perhaps most important of these is the concept that the effect of a particular stress at a particular time on the reproductive axis of an individual animal can be modulated by a great number of variables. These modulating variables include social status, the magnitude and duration of stress, perception of the stress, aggressive behavior, seasonal cues, and the prior level of activity within the reproductive axis (see **Figure 1**). Thus, although it is possible to measure group mean responses to specific stresses and make conclusions about the effect of a stress on the reproductive axis in that species, such a mean assessment may be of little use in determining whether an individual animal will experience a suppression of reproductive function in response to that stress. An increased understanding of the mechanisms by which psychological and social stresses suppress the activity of the reproductive axis may well be achieved by focusing future studies on these individual differences in response to stress.

See Also the Following Articles

Opioids; Primate Hierarchies and Personality; Social Status and Stress.

Further Reading

Abbott, D. H., O'Byrne, K. T., Sheffield, J. W., et al. (1989). Neuroendocrine suppression of LH secretion in subordinate female marmoset monkeys (*Callithrix jacchus*). In: Eley, R. M. (ed.) *Comparative reproduction in mammals and man*, pp. 63–67. Nairobi: National Museums of Kenya.

Adams, M. R., Kaplan, J. R. and Koritnik, D. R. (1985). Psychosocial influences on ovarian endocrine and ovulatory function in *Macaca fascicularis*. *Physiology and Behavior* **35**, 935–940.

Bowman, L. A., Dilley, S. R. and Keverne, E. B. (1978). Suppression of oestrogen-induced LH surges by social subordination in talapoin monkeys. *Nature* (*London*) **275**, 56–58.

French, J. A., Abbott, D. H. and Snowdon, C. T. (1984). The effect of social environment on estrogen excretion, scent marking, and sociosexual behavior in tamarins (*Saguinus oedipus*). *American Journal of Primatology* **6**, 155–167.

Keverne, E. B., Meller, R. E. and Eberhart, A. (1982). Dominance and subordination: concepts or physiological states? In: Chiarelli, O. (ed.) *Advanced views in primate biology*, pp. 81–94. Main Lectures of the VIIIth Congress of the International Primatalogical Society, Florence, 7–12 July 1980. New York: Springer-Verlag.

Nishida, T. (1983). Alpha status and agonistic alliance in wild chimpanzees (*Pan troglodytes schweinfuthii*). *Primates* **24**, 318–336.

Rowell, T. E. (1970). Baboon menstrual cycles affected by social environment. *Reproduction Fertility And Development* **21**, 133–141.

Sade, D. S., Cushing, K., Cushing, P., et al. (1976). Population dynamics in relation to social structure on Cayo Santiago. *Yearbook of Physical Anthropology* **20**, 253–262.

Sapolsky, R. M. (1982). The endocrine stress-response and social status in the wild baboon. *Hormones and Behavior* **16**, 279–292.

Wilson, M. E., Gordon, T. P. and Bernstein, I. S. (1978). Timing of births and reproductive success in rhesus monkey social groups. *Journal of Medical Primatology* **7**, 202–212.

Resistance

L M Zabarenko
Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by L M Zabarenko, volume 3, pp 373–375, © 2000, Elsevier Inc.

Introduction
Definitions
Implications for Technique
Implications for Theory
Current Views
Recent Research
Conclusion

Glossary

Intellectualization	The use of intellectual activities to exert control over anxiety and reduce tension. This proclivity is often encountered as a defense against instinctual drives in adolescence and is enlisted as resistance in insight-oriented therapy.
Psychoanalysis	A discipline discovered by Sigmund Freud which is (1) a psychotherapeutic technique, (2) a systematized theory of human behavior including development and (3) a method of investigating the mind.
Resistance	A familiar but paradoxical phenomenon in psychotherapy and psychoanalysis. The patient, having entered treatment, opposes the process in ways which can be conscious, pre-conscious or unconscious.
Transference	The displacement of patterns of feelings originally experienced in childhood onto a person involved in a current relationship.
Transference resistance	The usually unconscious use of transference affects as a resistance to therapeutic progress.

Introduction

By the time people consult a mental health professional for relief from the consequences of stress, they are usually highly motivated. Psychological suffering; anxiety and depression; and loss of functions such as the inability to work, engage fully in interpersonal relations, or care for themselves have often reached profound levels. It is remarkable, then, to discover that many patients also resist treatment and that such opposition is a reliable phenomenon that can be expressed in a whole universe of behaviors. The patient may be silent, withhold essential information from the therapist, flood the sessions with intellectualizations or dreams so that reporting consumes all the time, arrive late for appointments or fail to come, delay paying the fee or refuse to pay at all, and find it impossible to follow directions about medications. Many additional forms of resistance have been reported, including eagerness to make the best use of the time in the sessions, "talking over the treatment every day with some intimate friend," what Freud called a "therapeutic leak," using evasive speech (or lapsing into technical jargon) or clichés, or mumbling and accusing the therapist of suggesting ideas too strongly. There is also documentation of resistance as it occurs in the termination phases of psychoanalysis in children and in adults.

Patients can be aware of resistance, but often they are not; that is, resistance may be conscious, preconscious, unconscious, or a mix of all three. Regardless of the etiology, resistance tends to be most pronounced in the insight-oriented psychotherapies, those aimed not only at symptomatic relief but also at self-knowledge.

Definitions

In *The Interpretation of Dreams*, Freud was crystal clear: "Whatever disturbs the progress of the work is resistance"; this concept is still current. Stressing that resistance is ubiquitous and pervasive, Moore and Fine (1990: 168–169) also point out that "analysis of resistances ... has become a hallmark of psychoanalytic treatment."

Resistance as a Defense

In psychoanalysis, the most ambitious form of psychotherapy, patients are often unconsciously anxious about revealing themselves, dreading possible rejection by the analyst. They may also fear health itself as alien, dangerous, and possibly more menacing than being sick.

Under such circumstances, almost anything can be drafted to slow the process of change, as the lists described earlier attest. In addition, patients may treat the therapist as an enemy and alternately comply with and rebel against his or her presumed wishes; for example, some patients report memories or dreams but cannot work with them via associations. Among the feelings defended most keenly are those toward the therapist (i.e., transference resistances), but unconscious resistance may also be evoked in the analyst or therapist

and these constitute a manifestation of countertransference problems. These can interfere with the therapist's ongoing empathic resonance with the patient, resulting in technical misjudgments such as inappropriate content or timing of interpretations.

Resistance Resulting from Shifts in the Therapeutic Alliance and as Distinct from Negative Therapeutic Reaction and Negative Transference

Because it is pervasive, resistance should be clearly distinguished from other therapeutic phenomena, such as negative therapeutic reactions or negative transference. According to Moore and Fine (1990: 168–169), a negative therapeutic reaction is a "clinical response. . . . Following a period of what seems to be constructive and effective therapeutic technique and understanding, the patient's condition paradoxically worsens." Freud's position was that "Every partial solution that ought to result . . . in an improvement or a temporary suspension of symptoms produces . . . for the time being an exacerbation of (the) illness." However, when the patient sees the therapist as malevolent, exploitive, or manipulative, these are negative transference feelings.

History

Sigmund Freud is generally credited with the identification and explication of resistance in contemporary psychological thought. His first mention of the entity came in a discussion of Elizabeth von R., a patient who was sent to him in the fall of 1892 and became his "first full-length analysis of a hysteria." Instead of using hypnosis, as he had previously in his work with Josef Breuer, Freud encouraged this patient to associate freely, to say whatever came to mind. When she was silent and he inquired what was going on in her head, she replied, "Nothing." Freud refused to accept this answer, believing that it was the same species of willful forgetting that had caused her hysterical symptoms in the first place.

Another patient, the obsessional neurotic known as "the rat man," began treatment in late 1907. This man described the conflict that results in resistance by quoting Nietzsche to Freud: "'I did this,' says my Memory. 'I cannot have done this,' says my Pride and remains inexorable. In the end – Memory yields." Describing a kind of punishment practiced in Asia, the patient rose from the couch and begged Freud not to insist that he go on. Freud explained that "The overcoming of resistances is a law of the treatment" and helped him to finish the story; the perpetrator was tied down, a pot with rats in it was turned upside down on his buttocks, and the rats bored their way into his anus.

By 1912 Freud was confident enough to emphasize the pervasiveness of the phenomenon: "(It) accompanies the treatment step by step. Every single association, every act of the person . . . must reckon with the resistance . . . (and it) represents a compromise between the forces that are striving towards recovery and the opposing ones." Everything that distracts the patient from following the fundamental rule of free association is an obstruction.

Implications for Technique

Because resistance is a dramatic and undeniable example of the operation of the unconscious, its appearance can signal when therapeutic work is approaching painful and important areas of the illness. Such nodal points are often thick with cues about the etiology of the neurosis, and tact and timing are essential when attempting to bring the resistance to the patient's awareness. Careful and accurate interpretations at this point can inform both participants greatly and free up the therapeutic log jam.

Implications for Theory

At first thinking that resistance was simply an automatic defense, Freud saw it chiefly as an obstacle to the work. However, as the implications and extent of unconscious resistances became clearer, these led Freud to reconsider his topographic model of the psyche as the relationship of conscious, unconscious, and preconscious mentation and to incorporate these concepts into a new stance that stressed a structural view of the mind as composed of the id, ego, and superego.

One good way to grasp how profoundly the discoveries about resistance changed psychoanalytic theory is to contrast the taxonomies Freud used to categorize them. In 1895, he thought resistances appeared under three circumstances: (1) if "there is a personal estrangement . . . if the patient feels that she has been neglected, insulted, too little appreciated," (2) if the patient dreads becoming too dependent on the physician or influenced by him, or (3) if the patient is frightened at finding "she is transferring on to the figure of the physician the distressing ideas which arise from the content of the analysis." Coincidentally, this discussion marks the first use of the term transference in the psychoanalytic sense.

Twenty-eight years later in a famous passage, Freud found he required a five-tier classification system. Three species of resistance came from the ego – repression, resistance, transference resistance – and the unwillingness to renounce the gains of being ill was "an assimilation of the symptom into the ego."

Id resistance, approached by "working through" the analytic discoveries, and superego resistance completed the list. The latter, "the last to be discovered, seemed to spring from guilt or the need for punishment and opposes every move toward success."

Current Views

The understanding of resistance has remained a salient feature of psychotherapy and psychoanalysis and a prominent focus of psychoanalytic technique. However, with the epigenesis of post-Freudian theories, each school has proposed new ways of thinking about the etiology of resistance and offered corresponding technical recommendations for dealing with it.

Self-psychologists believe that resistance is evoked by failure of the therapist's empathic attunement and that such failure will mirror the narcissistic patient's early deprivation. Object relations theorists stress the extent to which resistance occurs when there has been insufficient recognition of the real as well as the unique transferential relationship between the therapist and the patient. Brenner's revised structural proposals see resistance mainly as a defense that is a compromise formation in the patient's conflict between wishing to recover and wishing to remain ill. Schafer feels that resistance can be traced to insufficient understanding of transferential–countertransferential currents as these occur in analytic work.

Recent Research

The explosive proliferation of research in the neurosciences and the deepening sophistication of findings in cognitive psychology have profound implications for clinicians facing resistance in their daily work. There are venerable journals in each area, but they are not notably accessible to the others and few acknowledge or attempt to integrate the contributions from all three areas. *Neuro-psychoanalysis* is one of the most scholarly and useful of a group of new journals whose aim is to do this. First published in 1999, it is staffed by large and distinguished editorial boards in neuroscience and psychoanalysis. The lively ongoing disagreements in this journal and the roster of established experts as authors testify to the relevance of neuroscience to analytic and dynamic psychotherapeutic work and simultaneously to the triangulation of cognitive psychology, neuroscience, and psychotherapy.

For example, recent research in cognitive psychology has underscored the importance of continuity in forming psychological constructs and offers an appealingly quantitative explanation of the fact that resistance that disrupts continuity (i.e., missing sessions or coming late) can delay or even disrupt therapeutic progress. The experimental demonstration that remembering itself can cause inhibition suggests that when patients' recovery of important memories falter and transferential constructs might be invoked, the therapist could also benefit from information from relevant cognitive principles that would enrich understanding the difficulty and inform technique.

Clinician–researcher–scholars capable of interdisciplinary synthesis are rare but not unknown. In a 2003 summary of his work, for example, Levin illustrates the importance of learning in psychotherapy and the links among learning, psychoanalysis (especially but not exclusively transference), and the neurosciences, an emphasis he has espoused for decades. Citing recent research findings, he suggests using these to augment existing ways of thinking about resistance – that resistance occurs because "patients defend themselves against more novelty than they can process at a given time" (Levin, 2003: 82), a cognitive and neural overload.

Conclusion

The importance of the patient's resistance to treatment has escalated as the disciplines of psychotherapy and psychoanalysis have matured and become theoretically and technically more diverse. However, the observation that there are obstacles to cure has remained consistent, as has the indefatigable work of finding more effective ways to understand and clear away these barriers to progress.

See Also the Following Article

Psychoanalysis.

Further Reading

Brenner, C. (1982). *The mind in conflict*. New York: International Universities Press.

Dewald, P. A. (1980). The handling of resistances in adult psychoanalysis. *International Journal of Psychoanalysis* **61**, 61–70.

Eder, M. (1930). Dreams as resistance. *International Journal of Psychoanalysis* **11**, 40–47.

Fraley, R. C. and Roberts, B. W. (2005). Continuity: a dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review* **112**, 60–74.

Freud, S. (2000). The interpretation of dreams (1900). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vols. 4–5), pp. 1–630. New York: W. W. Norton.

Freud, S. (2000). Notes upon a case of obsessional neurosis (1909). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 10), pp. 153–257. New York: W. W. Norton.
Freud, S. (2000). The dynamics of transference (1912). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 12), pp. 98–108. New York: W. W. Norton.
Freud, S. (2000). The ego and the id (1923). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 19), pp. 1–66. New York: W. W. Norton.
Freud, S. (2000). Inhibitions, symptoms and anxiety (1926). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud* (vol. 20), pp. 77–175. New York: W. W. Norton.
Gay, P. (1988). *Freud: a life for our time*. New York: W. W. Norton.
Kohut, H. (1971). *The analysis of the self*. Madison CT: International Universities Press.
Levin, F. (1985). The need for a psychoanalytic learning theory. Paper presented at the meeting of the American Society of Adolescence, Dallas, TX, May 17.
Levin, F. (2003). *Psyche and brain. The biology of talking cures*. Madison, NJ: International Universities Press.
Moore, B. and Fine, B. (eds.) (1990). *Psychoanalytic terms and concepts*. New Haven, CT: American Psychoanalytic Association and Yale University Press.
Schafer, R. (1997). Vicissitudes of remembering in the countertransference: fervent failure, colonisation and remembering otherwise. *International Journal of Psychoanalysis* 78, 1151–1163.
Stone, L. (1973). On resistance to the psychoanalytic process. In: Rubinstein, B. B. (ed.) *Psychoanalysis and contemporary science*, pp. 42–73, New York: Macmillan.
Veling, H. and van Knippemberg, A. (2004). Remembering can cause inhibition: retrieval-induced inhibition as a cue independent process. *Journal of Experimental Psychology* 30, 315–318.

Restraint Stress

R J Servatius
New Jersey Medical School and New Jersey Health Center, D. V. A. Medical Center, East Orange, NJ, USA
G Salameh and K M Coyle
New Jersey Medical School, Newark, NJ, USA
W P Paré
D. V. A. Medical Center, Perry Point, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 376–377, © 2000, Elsevier Inc.

Methods of Restraint
Stressor Intensity
Stressor Parameters

Glossary

Chronic stress state	A persistent presence of physiological and psychological symptoms of the stress response following exposure to an extreme stressor.
Inescapable stress	A stressor that is delivered to an animal whereby it cannot escape or avoid by performing any particular behavior.

Methods of Restraint

Loose Restraint

A method of restraining an animal so that free movement is restricted, yet some movement is still possible. Loose restraint is often used in conjunction with other procedures, not necessarily related to stress research, where it is necessary that the animal's movements be minimized, e.g., while taking a blood sample or administering an injection. In rodents, loose restraint is usually accomplished by placing the rodent in a small chamber or in a cloth. In nonhuman primates, restraint is accomplished by making the size of an individual cage smaller.

Tight Restraint

This method completely eliminates an animal's ability to move. Techniques for rodents include jacket or harness restraint or securing the rodent's limbs as it is lying prone. For nonhuman primates, tight restraint is accomplished by placing the subject in a chair with straps attached.

Supine Restraint

A variation of tight restraint for rodents in which the animal is tied down while lying on its back.

Combination Restraint Stress with Other Stressors

Restraint stress is combined with another type of stressor such as cold, heat, water immersion, and electrical shock to the tail.

Stressor Intensity

Hans Selye performed the pioneering work with restraint stress. The use of restraint, as an experimental procedure, was initially developed by Serge Bonfilo in France and by David Brodie in the United States. Subsequently, restraint stress has been utilized extensively, alone or in combination with other stressors, in order to measure its physiological, psychological, and pharmacological effects. Similar to pharmacological studies, the capacity of a particular treatment to produce a credible effect is reinforced by what is known as the "dose-response" relationship. The basic tenet of this relationship is that increasing doses of a particular treatment elicit increasing degrees of that treatment's effect. Therefore, based on integrated stress research spanning the fields of psychology, neuroscience, endocrinology, immunology, gastroenterology, cardiology, and pharmacology, one may grade a particular stress protocol's "dose" or intensity. In their two comprehensive reviews of the restraint stress literature, Paré and Glavin have provided the means to gauge the intensity of various experimental manipulations, including the parameters of restraint type, duration of restraint, and number of stress episodes, among others. Intensity is verified through associated effects on behavioral and physiological stress respondents. Thus, the methods of restraints can be graded as combination > supine > tight > loose.

Stressor Parameters

Duration of Restraint

While the type of stress most accurately predicts the intensity of a stress episode, the duration of the episode can greatly affect its intensity as well. In general, increasing the duration of a particular restraint session increases the degree of stress-associated symptoms. However, with the milder restraint types (loose restraint), longer duration bouts may be associated with decreased responsiveness—evidence of habituation.

Number of Restraint Episodes

As with the duration of a restraint session, increasing the number of sessions has the general effect of inducing more pathological consequences. However, this generalization is qualified by studies demonstrating diminished physiological responsiveness with multiple stressful episodes.

Other Factors

Several factors will influence the degree or duration of stress responses to restraint greatly.

1. Gender: pathophysiological consequences from restraint stress are more apparent when restraint stress episodes occur in the activity phase of estrus as opposed to the inactivity phase.
2. Cycles: the effects of restraint stress appear to differ as a function of time of day and season.
3. Age: although the amplitude of glucocorticoid responses in aged rats does not differ from young rats exposed to restraint stress, the duration of these elevations is longer in aged rats.
4. Strain: some rat strains (e.g., Wistar–Kyoto) are more vulnerable to the effects of restraint, whereas others are more resistant.

See Also the Following Articles

Animal Models (Nonprimate) for Human Stress; Immobilization Stress; Selye, Hans.

Further Reading

Glavin, G. B., Paré, W. P., Sandbak, T., Bakke, H. and Murison, R. (1994). Restraint stress in biomedical research: an update. *Neuroscience and Biobehavioral Reviews* **18**, 223–249.

Paré, W. P. and Glavin, G. B. (1986). Restraint stress in biomedical research: a review. *Neuroscience and Biobehavioral Reviews* **10**, 339–370.

Revenge Fantasies

M Horowitz and S Meffert
University of California, San Francisco, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Biological Issues
Social Issues
Psychological Issues
Treatment Issues
Conclusion

Revenge fantasies have been an important human activity since the beginning of recorded history. The Bible, Shakespeare, novels, and modern television and movies motivate their characters with fantasies and plans for seeking vengeance. The more traumatic an experience, the more intense are the revenge fantasies. Everyone is so familiar with revenge fantasies that there is no need to review their many forms. What is of concern is how to control revenge fantasies, how to reduce preoccupation with them, and how to prevent maladaptive action.

Biological Issues

The theme of revenge fantasies is so pervasive in human history, so common throughout populations that it seems to function almost like an instinct or drive. There is an adage that revenge is a dish best served cold. The indication of a biological function is not that revenge is cold; rather, it is a hot thirst that seems to be slaked when the revenge is taken. Animals other than human beings may take a revenge for a recent attack. Whenever a human being is caught up in a primal revenge action plan it is possible to use higher psychological processes to increase self-governance and to provide social organizational structures to prevent damage.

Social Issues

Genocidal campaigns are usually based on propaganda that incites bloodlust on the basis of revenge fantasies. General secular values are needed to counteract some fundamentalist religious values that indicate the need for blood revenge, as in killing whole populations. The evolution of society itself shows this history over several centuries. Social schemas of understanding and compassion have been used for this purpose. Our focus is mainly on psychological issues, so we comment here that, after a traumatic experience shared by a population group, it is important to emphasize to that population what is adaptive and maladaptive in terms of themes of retaliation versus those of adequate compensation or preparation so that the catastrophe will not happen again.

As recently noted by the World Health Organization, there is a divide in humanitarian aid services between mental and social health, particularly around the issue of trauma. Many debates have centered on the idea that posttraumatic stress disorder (PTSD) requires an understanding of individual pathology and that treatment that should be indexed to a particular culture, population, and era; critics argue that PTSD diagnosis and treatment are therefore not generalizable, particularly not across cultures. Such debates have become so heated that humanitarian aid efforts often have to choose between using individual, trauma-based mental health treatment and broader psychosocial public health-oriented treatment to address community-level mental health problems. The problem with splitting treatment into individual and social components is that it may obscure their heavily intertwined nature by compartmentalizing them into separate categories.

One of the most potentially destructive aspects of trauma is the impulse for retaliation that follows. The tragic aspect of this impulse is that it is much more often enacted against the close social contacts of the traumatized individual than the original aggressor; although the focus of revenge fantasies may be the hated enemy, the target is often loved ones, such as spouses and children. Research with Vietnam veterans shows that veterans with symptoms of psychological trauma are at high risk of committing violent acts toward their spouses and children.

One horrifying tactic of warfare capitalizes on this feature of human nature by raping enemy women or women who belong to the same social group as that of the enemy. This appears to be calculated not only to infect bloodlines; humiliate, torture, and sometimes kill the victim; but also to destroy social bonds in the enemy group. The rapists understand that the desire for revenge and anger against the group they represent will cause the community to reject the raped women from their social roles of wife, mother, and relative. The rapists count on the idea that feelings of anger and revenge will triumph over the social bonds of love and affection, leading to the internal destruction of the community. Unfortunately, they are often right.

Psychological Issues

Emily Dickinson may have led her cloistered, unmarried, and childless life because she felt betrayed. She became preoccupied with love and death in her poetry; in which she could only fantasize revenge. Here, for example, is her poem #339F.

I like a look of Agony,
[perhaps the look she would like to imagine on the face of this man]
Because I know it's true
[not false seductive expression, which she may have felt lured by]
Men do not sham Convulsion
[meaning the signs of suffering, although they may feign love]
Nor simulate, a Throe

The eyes glaze once – and that is Death –
[she is imagining as she writes that this man's death was painful, as he deserved]
Impossible to feign
The Beads upon the Forehead
By Homely Anguish Strung

One of the worst betrayals is when a trusted person exploits us. In modern days, some men insist on going bareback into intercourse (i.e., they refuse to use a condom). Some do so knowing that they risk transmitting a disease they know they have. The sexual partner finds out later that he or she has been infected with human immunodeficiency virus (HIV), hepatitis C, syphilis, gonorrhea, chlamydia, or genital herpes. Revenge fantasies then become a frequent compensation and sometimes a violent enactment.

The recipient of the disease may have felt love but that turns to hate. Revenge fantasies accumulate; forgiveness is beyond imagination. Still, the goal is to avoid violence or identifying with the selfish lover. Struggling to master a betrayal does not mean either taking revenge or giving up any course of compensatory action.

An effective response to a trauma is a course of action that may create more good than evil. Telling a community circle about the bareback perpetrator does so because it may prevent similar crimes. But figuring out the differences between a maladaptive hunger for revenge and an effective course of action is difficult. Expressions from the Bible such as "an eye for an eye, a tooth for a tooth" are uttered frequently. In our contemporary understanding of this language, they have a bloody, excessive connotation, which was not meant when the text was written. Then it meant do not annihilate or destroy the betrayer excessively but rather take a measured and reasoned course by retaliating by taking only an eye for an eye or a tooth for a tooth. In other words, do not massacre.

The perpetrator of a trauma that leads to revenge fantasies is seldom a total monster; we can try to understand his motives. Having some understanding does not mean blindly trusting the same abusing person or group again. Revising beliefs and values can lead to increased skill at seeing other humans as blends of good and evil, caring and selfishness. We get better at reading their intentions and learn when and how to be self-protective.

This work of reviewing memories of a trauma or betrayal can be painful. Intense negative emotions are rekindled. The strength of revenge fantasies may lead to repeated intrusive episodes, making it hard to concentrate our attention elsewhere. Our anger, fear that no one can be trusted, and despair over the coldness of the world can cause us to feel bitter or to impulsively engage in acts of hostility. Sometimes our aggression is displaced on to some weaker target. The crippling interpersonal effect of revenge fantasies is that, whether fulfilled or not, they often result in more destruction of those we love than of our despised enemies. Such venting leads to guilt and shame.

How can we avoid recycling trauma and revenge that causes further trauma in this manner? Thinking of a betrayal over and over again, without progress in healing; only increases bitterness and diminishes our chances at happiness. Our goal must be a point of completion in our processing of the memory. This may involve efforts to achieve a new more correct and realistic judgment. A plan for making a measured response at the right target or at no target can result.

A memory of an injury or an insult is fertile ground for growing a seed of rage into a towering weed that can strangle love and compassion. Some people or some groups always seem to carry a chip on their shoulder; they expect frustration from others and even nice gambits from others can be twisted and become misinterpreted as not enough. They feel as if the whole world has already betrayed them and refuses to make up for it. Each new betrayal reinforces this irrational attitude.

Seething with anger does have an energizing function because possessing targeted anger often makes us feel strong, whereas fear makes us feel weak. That is why states of self-righteous indignation are so popular on talk radio and why the revenge ideas are easy to sow and hard to give up.

After a betrayal, a person with such character traits of bitterness and resentment often enters a state of self-righteous indignation and then stays in this state because it feels like energy or fuel for the self. The burning anger helps such people feel more solid instead of feeling frail, empty, or apathetic. Breaking free of such snags is difficult. We can, however, learn to recognize the trap and step away from it. With the

recognition that we can be strong without being in a self-righteous rage, we can begin to analyze the situation rather than replaying an endless loop of revenge that goes nowhere.

A character trait of bitterness can result from unfortunate experiences in childhood or adolescence. Rightly or wrongly, many children regard their parents as omnipotent and so blame their parents for any suffering they felt as a consequence of unfortunate events. Any new frustration that occurs in their adult lives just adds to the trait of bitterness.

Treatment Issues

There are many evidence-based treatments for depression and anxiety, yet there remain very few treatments for the anger and interpersonal violence that often stem from trauma and revenge fantasies – with the exception of dialectical behavioral therapy, which generally aims to treat anger that results in self-harm. Violent behavior and the enactment of revenge fantasies often find their treatment, or lack thereof, in the criminal justice system, sometimes with court-mandated anger-management programs, for whose success there is little empirical support. Many factors may contribute to this situation, one of which could involve a conflicted attitude that society has toward anger, violence, and revenge. Particularly with cultural constructions of masculinity, the qualities of assertiveness, forcefulness, and even a measure of anger or the implication of violence are sometimes valued as reflecting ambition, competence, or efficiency in achieving life goals. Hollywood, of course, is a frequent purveyor of these stereotypes, with the hot-tempered violent hero with a chip on his shoulder from some previous wrong. These concepts stand in contrast to other cultural norms, such as the many aphorisms and maxims that underscore the destructive effects of anger and revenge.

As an adult, a person can realize the truth of parental limitations. These limitations, seen through adult mindfulness, are usually different from adolescent visions of how bad the parents had been. As a child, the person regards a parental lapse as being done on purpose. As an adolescent, he or she may show contempt for grownups. As an adult, the person can revise his or her entire internalized story of what happened and the explanations for how and why it happened. For example, a mother's depression and substance abuse can be seen as a terrible difficulty and struggle rather than as a deliberate abandonment of a child.

The story of a person's own life can be revised in the direction of having more sympathy for, and less blame toward, parental figures. This does not mean memories and attitudes toward parents are changed to happy ones. Many adults had unerasable childhood tragedies and or deficiencies in parenting. On adult reflection, there may be a realistic appraisal of the past events, but that realistic version may still fall far short of some ideal or even good-enough parenting.

Realistic appraisal may lead to sadness, even a delayed mourning for childhood losses. The person must mourn a childhood that cannot ever be repeated with a healthier-minded parent. Passing through that grief makes the person ready to seize what the here and now might provide and to look to the near future for what is possible. New, warm, loving relationships in various areas of one's life can help one to heal the wounds of childhood. The person learns to forgive him- or herself because the person was not to blame, as he or she might have supposed as a child, and to forgive others for the neglect they could not prevent.

Suppose the person had an abusive and neglectful father, became bitter, and now unconsciously hoped to find an ideal father as an adult. The person hopes that new candidates will make up for the damage done in the past, but they all turn out to fail the mission. At first, the new candidate for caring seems amiable and supportive, but the good qualities fall far short of the person's immense expectations. Sometimes others are provoked to fail; an otherwise inexplicable hatred is then felt toward the contemporary nice person and the inappropriate anger ruins the relationship. Efforts to change can stop this vicious cycle and reduce even the character trait of being constantly bitter.

The general wisdom of pursuing forgiveness does not condone neglect or abuse; it simply understands the humanness in all parties. The person avoids demonizing the targets of his or her moral indignation. Compassion is sought, but this does not mean passive resignation to enduring wrongs. Turning the other cheek may only invite more slaps in the face.

Giving up self-righteousness is not the same thing as giving up moral indignation. Moral indignation is a response to seeing people break rules when they could follow them. A good person can use moral indignation and do something to prevent evil and promote compassion or virtue. He or she may use moral indignation to gather a like-minded group together to encounter and counteract the exposure of children to abuse or neglect. This kind of indignation is effective and adaptive, whereas chronic bitterness with revenge fantasies is inefficient and maladaptive.

After experiencing an intense trauma or betrayal, the person has to ask him- or herself whether the perpetrator is a friend, an enemy, or perhaps an unreliable narcissist. The person modifies his or her expectations according to a realistic appraisal. The

person must learn to recognize enemies only when they exist. Even in these situations, the best aim is to temper wrath with reason because the sword of destructiveness often cuts backward.

It is helpful to ask questions. Why did the enemy attack or deceive? How was I harmed? If I counterattack, what are the likely consequences? How does a cycle of betrayal, attack, and counterattack end? In answering these questions, the person finds that the extreme of an ideal self and a demonic enemy is almost always wrong. The person must clarify more realistic central appraisals.

Suppose there is a persistent enemy. One ideal scenario is to trust in a truce based on a pact of nonbelligerence. Can a person realistically find a way so that the enemy will respect the rules? In the dreaded train of thought, the enemy is divested of human qualities and is viewed as a constantly malignant force. In an ideal train of thought, the enemy keeps his or her word of honor. These opposing extremes set the stage for a moderated train of thought. The rational analysis does not automatically arrive at truce or warfare. The enemy may remain the enemy. The rational appraisal leads to preparation for defense and continued alertness to signs of increasing danger. It rejects an attitude of unending battle and vendettas. It sets up bad consequences for breaking a truce and good consequences for keeping it. The person may fight the enemy but plan to accept the enemy's child.

Conclusion

The best adage ever coined against brutality is the golden rule: do unto others as you would want them to do unto you; do not treat others as you would not want to be treated. An excellent quality of such adages is that they can become a socially shared morality or ethic that can be taught by any religion or culture.

A new adage might be: when you feel bitter, seek the good. This adds something to existing popular adages such as: perform random acts of kindness and senseless acts of beauty. Usually people want to do good when they feel good or are in love. The goal is to do good when we are less elated. The idea of bringing goodness into the world when we feel demeaned, insulted, angry, and frustrated is a useful one to highlight because this stance can counteract self-righteous rage and revenge impulses.

This attitude – when you feel bitter, do good – counteracts two normal but malignant responses to being insulted. One response is to cave in and feel weak, deflated, and degraded; the other is the opposite, a result of role reversal, to adopt a strong menacing role from which to lash out at others who are weaker. For example, a child who sees a civil war squad murder his parents may well grow up to become a civil war soldier ready to commit similar atrocities or worse.

Adages help keep us alert, so here are some other versions. When you feel bitter, do good. Do not wait until you feel good to act with compassion for others. Protect others as you want to be protected. The best cure for an enraged heart is comforting others. Do not pass hatred on; fight it with compassion.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Disasters and Mass Violence, Public, Effects of; Domestic Violence; Male Partner Violence; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of; Violence; War Stress in the Former Yugoslavia; War-Related Posttraumatic Stress Disorder, Treatment of; War, Suicide and Sacrifice; Posttraumatic Stress Disorder – Clinical; School Violence and Bullying.

Further Reading

Cardozo, B. L., Kaiser, R., Gotway, C., et al. (2003). Mental health, social functioning, and feelings of hatred and revenge of Kosovar Albanians one year after the war in Kosovo. *Journal of Traumatic Stress* **16**, 351–360.

Dickenson, E. (1924). *The Complete Poems.* Boston: Little, Brown.

Horowitz, M. J. (1998). *Cognitive psychodynamics: from conflict to character.* New York: John Wiley.

Horowitz, M. (2001). *Stress response syndromes* (4th edn.). Northvale, NJ: Aronson.

Horowitz, M. (2004). *Treatment of stress response syndromes.* Washington, DC: American Psychiatric Press.

Horowitz, M. J. (2005). *Understanding psychotherapy change.* Washington, DC: American Psychological Association.

Orth, V., Montad, L. and Maercker, A. (2006). Feelings of revenge, retaliation motive, and posttraumatic stress reactions in crime victims. *Journal of Interpersonal Violence* **21**, 229–243.

van Ommeren, M., Saxena, S. and Saraceno, B. (2005). Mental and social health during and after acute emergencies: emerging consensus? *Bulletin of the World Health Organization* **83**(1), 71–77.

Young, A. (1997). *The harmony of illusions: inventing post-traumatic stress disorder.* Princeton, NJ: Princeton University Press.

Rheumatic Disorders

A T Masi and J C Aldag
University of Illinois College of Medicine at Peoria (UICOMP), Peoria, IL, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction
Prototype Rheumatic Disorders and Stressors
Possible Stressor Specificities to Particular Pathways
Profoundly Complex Patterns of Stress Responses
Stressors Related to the Physiopathogenesis of Rheumatoid Arthritis
An Evolving Research Challenge in Rheumatoid Arthritis
Summary

Glossary

Ankylosing spondylitis A form of arthritis mainly affecting the spine and sacroiliac joints and, to a lesser degree, those of the girdles and lower extremities. A striking association occurs with the genetic factor HLA-B27 and familial aggregation is strong. Mechanisms remain obscure, but typical enthesopathic lesions suggest that biomechanical stresses could activate inflammatory pathways in its physiopathogenesis.

Biomechanics The study of the application of mechanical laws and the actions of internal or external forces upon the living body or its structures. Biomechanical forces are believed to be important contributors to a number of rheumatic disorders, especially ankylosing spondylitis and osteoarthritis.

Central sensitization A biopsychological process whereby sensory signals or stimuli are amplified in intensity with expansion of the receptive fields through central nervous system (CNS) mechanisms.

Central sensitivity syndromes Chronic, noninflammatory rheumatic disorders believed to constitute an interrelated or overlapping family of symptomatic and dysfunctional conditions. The causes are unknown, but may result from central sensitization mechanisms and may be amplified by environmental, personal, or socioeconomic stresses.

Enthesopathy A pathological process occurring at sites of attachment or insertion of joint capsules, ligaments, or muscle tendons into bony tissues, which is typically found in ankylosing spondylitis and related spondyloarthropathy conditions.

Fibromyalgia syndrome A common and prototypic example of the family of central sensitivity syndromes. It is defined by widespread pain and the physical examination finding of multiple foci of localized tenderness on firm palpation at characteristic points (tender points), particularly of myofascial tissues. Pain, stiffness, and systemic symptoms of poor sleep and fatigue, anxiety or depression are often aggravated by insufficient or excessive physical activity and by environmental, personal, or socioeconomic stresses.

Microvascular endothelium The single cell lining layer of the luminal surface of the microvascular system, which regulates transport of macromolecules and blood components into the interstitium. Endothelial cells are regulated by the neuroendocrine and immune (NEI) systems and contribute to inflammation by their activation leading to enhancement of inflammatory cell migration from the lumen into the interstitium, and by their secretion of mediators of inflammation.

Osteoarthritis The most common form of arthritis, particularly in older ages, which results from biomechanical stresses and biochemical changes that damage articular cartilage and also cause lesions in the surrounding osteoligamentous tissues of the joint. The process may be either primary resulting from chronic excessive repetitive joint impacting or secondary to gross trauma or inflammatory arthritis conditions. Weight-bearing joints are mainly affected and it is often referred to as degenerative joint disease.

Physiopathogenesis A process which progresses from chronic disordered physiology or dysregulated function of organs or systems (presymptomatic phase) to later stages of pathological, symptomatic disease states. The implication of this term is a transition from initial chronic physiological dysregulations (preclinical phase) to later clinical disease states. Alternatively, the term, pathophysiology, implies that initial pathological processes alter physiological functions secondarily. Bidirectional derangements often occur in many diseases, but to differing degrees, and depending upon the phases of the processes.

Rheumatoid arthritis A chronic systemic disease primarily affecting small joints of the extremities in a symmetrical fashion and characterized

by inflammatory changes in the synovial membranes. Without effective therapy, the pathological process extends beyond the synovial tissue and inflammation erodes adjacent joint cartilage and bone, leading to deformities and disability. The presymptomatic causes of this disease leading to inflammatory pathways are unknown, although such pathology predominates in the clinical phase. Women are affected about three times as often as men, which is opposite to ankylosing spondylitis, and is an important clue to respective predisposing mechanisms.

Rheumatic disorders — The general nosological term for conditions affecting the musculoskeletal system that include symptoms of pain, stiffness, and other related abnormalities pertaining to joints or nonarticular (myofascial) structures. These disorders include a wide variety of conditions characterized by degeneration, inflammation, or metabolic derangements of connective tissues, joints, or other musculoskeletal structures

Spondyloarthropathy disorders — A group of overlapping arthritic diseases affecting extremity or sacroiliac joints and the vertebral spine to varying extents. Ankylosing spondylitis is the prototypic disorder in this group that mainly affects the sacroiliac joints and the spine, but can also involve lower extremity and girdle joints. They share clinical and radiological features, particularly enthesopathy lesions and increased presence of the cell surface HLA-B27 antigen marker.

Introduction

Historical Contributions to Stress Concepts

Modern concepts of regulatory biology may be traced to Claude Bernard (1813–78), who pioneered the doctrine of a constancy of the *milieu interieur*. Walter Cannon (1871–1945) further promoted and broadened the concept of physiologic organization and regulations, especially in his 1929 report on homeostasis. That original concept referred to an ideal setting of effectors controlling the *milieu interieur*. However, modern concepts of homeostasis infer the occurrence of different settings of the effectors as well as the accumulation of stress impacts or allostatic load on the physiological control mechanisms. The paradigm of allostatic load proposes that stress mediators have protective as well as damaging effects on the organism. The accumulation and overexposure of the body to stresses can promote adverse effects on various systems by the mediators.

In the next decade, Hans Selye (1907–82) developed experimental (rodent) models of generalized stress reactions to noxious agents. He demonstrated opposite changes in the thymus (atrophy) and adrenals (hypertrophy) as part of the organism's responses to injuries and chemical intoxications. Those experiments revealed counterregulatory relations between the hypertrophy of adrenal glands and involution of lymphatic tissues, which he originally considered to be nonspecific.

In 1946, Selye published his synthetic theory of the general adaptation syndrome and diseases of adaptation. Relative adrenal insufficiency was proposed to be a contributory factor in stress-related or immunologically mediated diseases. Selye's research and that of others indicated that the adrenal glands are closely connected with the organism's resistance to noxious stimuli or stresses.

In humans, adrenal gland enlargement was recognized to result from a wide variety of conditions, including those of increased physiological demands (pregnancy, cold exposure, exercise, and low oxygen tension), hypermetabolic disturbances (hyperthyroidism, and excessive insulin administration), infections, burns, and other shock-producing states as well as multiple toxic drugs. The concept of stress reactions became recognized largely under Selye's influence. His theory offered a unifying mechanism to interpret complex sets of metabolic, psychoneurophysiological, immunological, and other adaptive changes that characterized responses to a wide variety of noxious or perturbing stimuli.

The above-mentioned and other concepts of stress biology will be related to selected examples of rheumatic disorders. Human stress responses are profoundly complex, but their diverse impacts may be inferred to contribute to symptoms, physiopathogenesis, or outcomes of particular rheumatic disorders.

A Broad Spectrum of Rheumatic Disorders

Rheumatic disorders encompass a broad spectrum of conditions which affect the musculoligamentous (myofascial) and skeletal structures of the body as well as other soft connective tissues and internal organ control systems. These conditions include over 100 medical diagnoses. The diseases usually have defining clinical features, which vary in patterns of pain, psychological behaviors, limitations of movements, localizations of inflammation, or structural involvements in the body. The patterns of symptoms or pathology result from varied, but often overlapping primary or secondary pathways in the causations of these diseases. Classification of these different disorders is often based upon defining symptoms or clinical patterns, rather than upon specific etiology.

Nosology of Rheumatic Disorders and Potential Stressors

The term, rheumatism, is derived from the Greek *rheuma*, flux, inferring that symptoms or manifestations can *flow* from one part of the body to another. The expanded classification of *Arthritis and Rheumatic Diseases* is the overriding medical label for this broad spectrum of conditions and diseases. The role of stress in such a large number and diverse patterns of rheumatic disorders cannot be fully addressed in this limited review. Particular examples and principles of stress reactions affecting this family of conditions are offered. Research indicates that the general population is often exposed to physical or psychological stressors, which vary considerably among individuals and subgroups. The patterns of stressors tend to relate to particular rheumatic disorders, presumably via their respective mechanistic pathways. Examples of types of stressors affecting rheumatic diseases are provided.

Prototype Rheumatic Disorders and Stressors

Features of Fibromyalgia Syndrome

A type of noninflammatory rheumatic condition associated with stress mechanisms is fibromyalgia syndrome (FMS). This disorder is manifested by musculoligamentous (myofascial) pain, fatigue, nonrestorative sleep, affective features, and decreased functional ability. This condition is an identifiable clinical entity, despite the absence of objective physical deformity, inflammation, or structural lesions. Its clinical symptoms follow rather characteristic patterns. Its course tends to be associated with the degrees of subjectively perceived stresses or objectively experienced physical aggravating factors. The causal pathways of FMS are unknown, but are believed to involve central nervous system (CNS) and autonomic nervous system (ANS) activation and sensitization. The more prominently implicated stressors are believed to operate via these physiological pathways.

The particular stress factors which aggravate FMS differ in individuals, perhaps due to their predispositions and personalized circumstances. They include variously described stresses in personal relationships, demands of daily living as well as particular environmental or occupational hardships. Additional individualized stressors include cold or unusually warm temperatures, trauma, physical impacting or other biopsychological stresses, and biomechanical strains on the body. Data suggest that sufficient intensity of stress can cause central sensitization, as reflected in posttraumatic stress disorder (PTSD).

FMS is believed to be but one recognized clinical entity in a family of central sensitivity syndromes. These disorders compose a spectrum of clinically overlapping and co-occurring noninflammatory conditions. They frequently manifest with myofascial pain, fatigue, nonrestorative sleep, hypersensitivity to various stimuli, and affective features. The conditions are somewhat diverse, involving different systems, but tend to have overlapping physiological stressors and aggravating environmental or occupational factors. Criteria have been developed to classify the particular syndromes, based upon relatively defining clinical features.

Features of Rheumatoid Arthritis

The main inflammatory rheumatic condition believed to be affected by stress mechanisms is rheumatoid arthritis (RA). This disease is characterized clinically by pain, objective swelling, and inflammatory damage in multiple extremity joints (i.e., polyarticular) in a symmetrical pattern. Blood samples of RA patients clearly show evidence of systemic inflammation. Increased levels of biomarkers (e.g., acute phase proteins) and mediators of inflammation (e.g., inflammatory cytokines) are found in the blood and joints during active phases of this disease.

In genetically susceptible people, the underlying systemic mechanisms believed to be operating in RA are dysregulations of the neuroendocrine and immune (NEI) as well as the microvascular endothelial (MVE) systems and their interactions. The more specific stressors are believed to further perturb the already dysregulated NEI and MVE systems or place increased demands upon joints. However, CNS-mediated psychosocial stresses also affect pain behaviors and disease activity in RA, as may occur generally in other disorders.

Stressors of the RA clinical status include particular physical (e.g., overuse of joints) or subjective (e.g., poor sleep and fatigue) factors as well as systemic circumstances, which overtax the physiological balance or energy reserves of the body. Aggravating or stressor factors can increase both symptoms and objective manifestations of RA. Such stressors are believed to operate via enhanced perturbations of the already dysregulated physiopathogenetic pathways in RA as well as by greater direct physical demands upon compromised joints.

Features of Osteoarthritis

Various biomechanical stressors operate in rheumatic disorders, to cause either aggravated symptomatology or accelerated tissue damage. Osteoarthritis (OA) and ankylosing spondylitis (AS) are diseases, which result in damage and deformity to peripheral or spinal joints. The mechanisms are believed to be primarily

caused by biomechanical stressors, at least in their initial stages. Later phases of these diseases may include some degrees of inflammation, either localized articular (OA) or combined joint and systemic (AS), but to lesser degrees than in RA.

Structurally, the basic elements of body design are balanced systems of compression and tension in its functioning components. Biomechanical stresses which alter such balance, e.g., excessive compression (chronic overloading from obesity) or excessive tension (suspected axial myofascial hypertonicity in AS) can contribute to accelerated joint damage or increased symptomatology.

In diarthrodial (freely moving) joints, movements occur across the smooth specialized cartilage tissue surfaces covering the ends of connecting bones (e.g., knuckles of the hands, hips, or knees). This remarkable cartilage tissue provides combined compressional elasticity (i.e., water trapping) and tensile strength (i.e., resistance to tearing). Although a popular lay concept is that joints are primarily designed to absorb impacts, their main function is to permit smooth movements in transmission of dynamic forces between connecting bones, ligaments, and muscles. Friction and impacting of joints are normally minimal, since surface movements (velocity vectors) are parallel to the joint contours, which have covering fluid films. The coefficient of friction across normally articulating joint surfaces is miniscule (Figure 1).

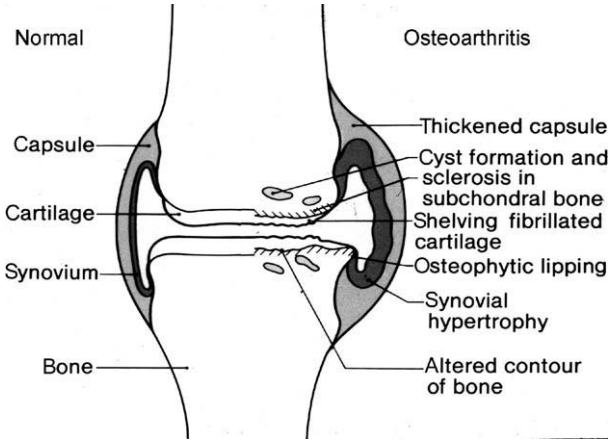


Figure 1 Schema of normal versus osteoarthritis joint. Anatomical features of a normal diarthrodial joint are schematically shown. The cartilage is a specialized tissue which normally provides compressible elasticity and resistance to tearing in smooth joint movements. Muscles and ligaments are attached to the joint capsule and bones, providing essential dynamic support and transfers of forces. In osteoarthritis, cartilage undergoes loss of tissue and function and the articulating bones undergo gross alterations in contours as well as defects in microarchitecture. The synovial lining of the joint becomes secondarily inflamed from reactions to the cartilage breakdown products and increased strains of a biomechanically stressed joint. Synovial changes in osteoarthritis are less manifest than in rheumatoid arthritis, where inflammation is a primary pathological process in the joint.

When joints are chronically overimpacted or overloaded, as in OA, the cartilage tissues begin to fail. Cartilage loses its normal elasticity, proper opposing surface conformations, and tensile integrity. The later stages of OA are characterized by substantial losses of cartilage substance, deformities of bones and joints (Figure 1). The consequences are loss of mechanical joint efficiency, pain on movements, secondary muscle atrophy, and decreased dynamic support. Thus, chronically decompensated or unbalanced biomechanical stresses can contribute to joint damage and later failure. Gross trauma can cause acute injury and loss of joint integrity, leading to subsequent chronic joint failure, as described above.

Features of Ankylosing Spondylitis

In AS, lesions usually begin in the sacroiliac joints (SIJs) or the lumbosacral vertebrae, at the base of the spine, where the truncal compressional forces are greatest (Figure 2). These joints have minimal, but essential, normal degrees of motion. They are primarily designed to provide stability as well as transfer of forces between the trunk and lower extremities and vice versa. The vertebral lesions in AS are mainly localized at the attachment tissues (enthesopathy), which result in secondary calcifications and immobilization by bony fusion of the connecting vertebrae. Typically, the spinal lesions progress from the base to higher levels, as limitations in movements and fixation occur below. In the SIJs, the pathological lesions progress from initial cartilage narrowing, to localized erosions of this tissue, and to final total cartilage loss, with obliteration of the joints by an intra-articular bony fusion process.

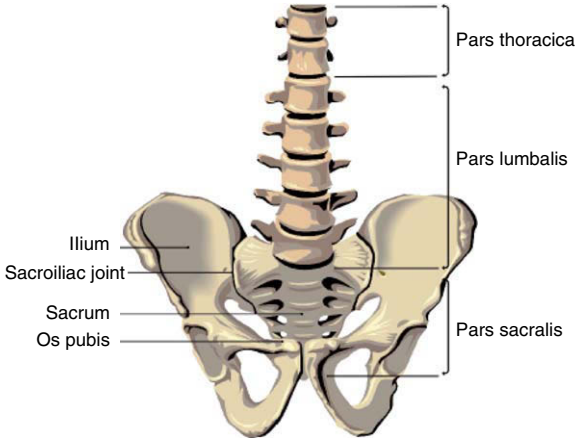


Figure 2 Skeletal structures relevant to ankylosing spondylitis. Skeletal structures of the lower vertebral spine, sacrum, and pelvis are shown. In ankylosing spondylitis, pathological changes typically occur first in the sacroiliac joints and the lumbosacral spine, which bear the greatest truncal loads. Biomechanical stresses are believed to play a major role in development and localization of tissue changes in this disease (see text).

Altered biomechanics is suspected to be important, if not primary, in the physiopathological processes in AS. One concept favored by these authors is that the SIJ lesions develop in a setting of excessive compressional forces across these joints, whereas the vertebral lesions result from excessive tensional stresses on the attachment tissues. In contrast, the popular concept is that inflammation is the primary mechanism at these sites. However, neither the initiating mechanisms nor explanation for the specific tissue localizations are elucidated in this inflammatory theory. If the above-mentioned biomechanical alterations are the main predisposing factors in AS, secondary inflammatory mechanisms could logically occur following chronic microinjury.

All organisms on this planet have evolved and adapted appropriately to resist forces of gravity. Accordingly, all normal tissues and body parts must possess sufficient cellular and structural integrity to resist gravity as well as to maintain an appropriate balance of stability and flexibility mechanisms to perform the vast spectrum of life's daily activities. Structural integrity of the musculoskeletal system is required at micro- and macroanatomical levels. Also, efficient biomechanical connectivity of the myofascial and osteoligamentous tissues is essential. Efficient transfers of physical energies and forces requires dynamically normal organization in the performance of the multitude of body movements and resistance activities. Compromised integrity of either the musculoskeletal structures or their dynamic organization predisposes to particular diseases like OA and AS as well as to accelerated joint damage in RA and other rheumatic conditions.

Possible Stressor Specificities to Particular Pathways

In order to better interpret how stressors can affect particular rheumatic disorders, it is important to understand the main mechanisms or systems alterations, which determine these respective conditions. For example, CNS pathways are believed to be the main mechanisms operating in FMS. They include enhanced emotional and sensory perceptual activation via autonomic and cognitive systems, rather than by alternative inflammatory or biomechanical stress mechanisms. Instead, dysregulations or activations of the NEI and MVE systems are believed to be the primary pathways in RA, leading to enhanced immunological and inflammatory reactivity. Whereas, biomechanical stressors are believed to cause the primary structural and tissue lesions in OA and AS.

Types of stressors that typically aggravate a particular rheumatic disorder may be related to the host's susceptibly to the respective condition. Stressor specificity relationships to rheumatic disorders might be metaphorically described as a particular glove fitting the owner's hand, at least in part. If stress-disorder specificity does occur, the phenomenon may reflect a potential biological band-width resonance between the stressors and reactor systems (as suggested by Dr. Richard Imrich). Future investigations may permit a partitioning of stresses and their reactive effects according to degrees of stressor intensity and frequency and categories, in relation to the host's adaptabilities versus vulnerabilities, under particular environmental and psychosocial circumstances. For example, experimental and human data suggest that repeated stress exposures under controlled and supportive conditions tend to increase resistance to adverse reactions (favorable conditioning) and vice versa (enhanced vulnerability).

Although certain stressors tend to primarily aggravate one or another rheumatic disease pathway, they can often indirectly or secondarily affect other disease systems. Thus, particular types of stressors are common in related rheumatic disorders, but they are not strictly compartmentalized in their effects. Stressors can have varied degrees of overlapping primary or secondary influences due to complex interactions of the host and environmental circumstances in a particular disorder.

An extensive spectrum of psychological, systemic, physical, or biomechanical loading factors can influence rheumatic disorders. Each condition tends to have its respective pattern of host susceptibilities, somatic developmental influences, psychosocial predispositions, or correlated behaviors. The different disorders show relative reactivity to the varied environmental stimuli, stresses, or challenges. Individuals differ in multiple and complex ways, and are affected differently by the respective heterogeneous stressors. Therefore, only general interpretations can be made regarding the main types of stressors affecting particular rheumatic disorders, and considerably more research is needed in this area.

Profoundly Complex Patterns of Stress Responses

Central Stress Systems

Different stressors of sufficient intensity can elicit relatively patterned responses via multiple effectors. The intensity of stressors and their consequent responses may be either relatively limited or more greatly enhanced. Limited psychophysiological stresses may primarily stimulate CNS and ANS pathways. Enhanced systemic stressors may also significantly activate the hypothalamic-pituitary-adrenal (HPA) axis and related neuroendocrine effectors. These include the two adrenal systems, i.e., the adrenocortical

(corticosteroids) and the adrenomedullary (adrenaline). Other HPA axis components contribute to stress-related responses, which are intended to diminish acute tissue injuries and modulate overstimulation of the immunological and inflammatory responses.

Neurodynamic models of the brain propose that sensory circuits can be tuned, for purposes of selectively sensitizing the cortices to desired or expected inputs. Such dynamics enhance selective attention and intentional learning. However, selective sensitivity and central sensitization in people susceptible to FMS and related conditions could possibly overamplify signals perceived to be averse. In overly distressed people, such mechanisms would need to be modulated or unlearned, in order to relieve sensory and psychic pain induced by stressful psychosocial stimuli.

Although ANS dysfunction may explain many of the multisystem features of FMS, a fundamental dilemma is whether such activated status is constitutionally predisposed or developmentally acquired. Some data indicate that HPA or ANS dysfunctions may be acquired responses in particular people (possibly innately susceptible) who are exposed to various stressful psychosocial or other life circumstances, particularly during childhood or developmental stages. Carefully controlled prospective studies in family settings will be needed to better assess the dynamic impacts of stress mechanisms on familial versus environmental backgrounds, and their interactions. Such data can clarify the respective roles of these components in FMS, related disorders, and symptomatic patterns of illnesses generally.

Responses to major emotional or physical stresses are physiologically integrated by the HPA axis, the adrenal medulla, and the sympathetic nervous system (SNS). People differ in their susceptibility to physical stresses or psychoneuroendocrine perturbations. Some people have limited versus robust stress reactions, depending upon the integrity of their physiological adaptation. Healthy, well-conditioned, younger people generally react more favorably to systemic stresses (i.e., less allostatic load), than elderly or frail individuals, who may more likely show adverse reactions and consequences.

Variability in Central Nervous System Stress Responses

Individuals vary in cognitive or perceptual sensory sensitization and in their SNS activation, due to genetic, developmental, and psychosocial or behavioral conditioning. Stresses that perturb these CNS pathways can amplify many subjective complaints by central sensitizing mechanisms. Such symptoms include greater intensity and distribution of pain, anxiety, or dysphoria, increased arousal and nonrestorative sleep, easy fatigue, and generally diminished functional ability. More limited intensities of psychosocial, physical, and environmental stresses may not be sufficient to perturb the majority of individuals in the population, nor be significant factors to predictably alter NEI–MVE interactions in the majority of people. However, people more susceptible to the preceding limited stressors (i.e., those having low stress tolerance) may nevertheless experience significantly enhanced symptomatology or dysfunctions.

In FMS patients, stress susceptibility pathways seem to be related mainly to increased central sensory sensitization mechanisms. Such pathways can enhance perceptual pain and autonomic emotive reactions, but not necessarily affect systemic neuroendocrine or immune mechanisms that predispose to inflammation. Susceptibilities to particular rheumatic disorders seem to be amplified by complementary stressors. In FMS, the more evident stressors seem to primarily impact central sensitization pathways. However, various stressors can secondarily influence sleep and physical conditioning, and can affect other body systems. In addition, emotional or biomechanical factors can also amplify symptomatology in FMS and related disorders.

Stress Reactivity and Systemic Pathways

In subsets of RA patients, stressors are believed to impact the HPA axis and MVE pathways. These systems may possibly show greater stress responsiveness in RA patients due to: (1) innate deficiencies in counterregulatory competence; (2) behavioral exposures or constitutional susceptibility to endothelial microvascular activation, or (3) constitutionally increased immunological or inflammatory activation, which are described further below.

Biomechanical Stresses and Anatomical Structures

Although, OA and AS are strongly affected by biomechanical stresses and gravitational forces, other rheumatic conditions can also be influenced by such factors. In the spinal column, vertebral bodies act mainly as compressional and supportive members, whereas the intervertebral anterior longitudinal ligament acts as a binding or tensional element. In the spine, forces are transmitted at specialized junctional sites, where spinal ligaments attach to the vertebral bodies (i.e., entheses). Such anatomical design allows both stability and mobility via the shared functions of the compressional and tensional osteoligamentous elements. However, such passive osteoligamentous structures alone do not permit sufficient stability to maintain anatomical equilibrium against gravitational forces, let alone provide for stabilized movements

or resistance activities. It is the skeletal and particularly postural muscles that provide the requisite supportive and movement functions in the body via their active contractile mechanisms and mechanical tone.

Complex Integrated Homeostasis and Stress Responses

Complex net interactions occur among the different homeostatic systems and subsystems. In physiological balance, prompt compensation occurs among the systems in response to varied changes. A primary versus secondary perturbation is difficult to distinguish, unless the pathways are known to be predetermined. In chronic disease, e.g., RA, dysregulations in the NEI and MVE systems occur concurrently. Their sequences of perturbations are not known and cannot be determined in cross-sectional analyses. The respective sequential changes must be analyzed over the course of time in prospective studies. Such longitudinal studies can better differentiate initial or primary alterations from secondary pathways, and are a priority objective for future research. Integrative longitudinal analyses offer greater promise to uncover complex mechanisms in chronic diseases with multifactorial determinants, than reductionist analyses which mainly address single-factor pathogenic mechanisms.

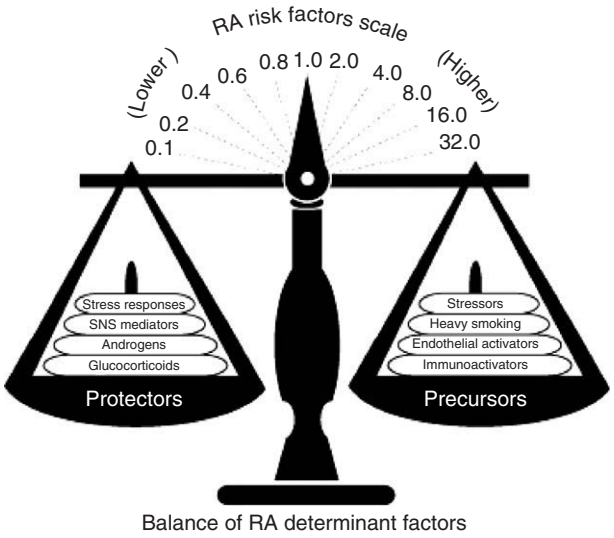


Figure 3 Balance of rheumatoid arthritis determinant factors. The recognized precursor risk factors and counteropposing protective factors for onset of rheumatoid arthritis are schematically shown. Normally, counterregulation occurs between activation of the immunological and microvascular endothelial systems and the anti-inflammatory systems, including glucocorticoids, androgens, and sympathetic nervous system (SNS) mediators. In RA, heavy cigarette smoking and other stressors are believed to dysregulate normal physiological balance and enhance activation of inflammatory pathways (see text).

Stressors Related to the Physiopathogenesis of Rheumatoid Arthritis

The major example of a rheumatic disorder in this article is RA. It illustrates important principles pertaining to the group of systemic inflammatory diseases and how they may be influenced by various stressors. Considering only the adult onsets of RA in this discussion, the clinical spectrum is decidedly heterogeneous in patterns of manifestations and in degrees of severity. Such disease heterogeneity (phenotypes) may be attributable to a combination of innate genetic predispositions, somatic developmental influences as well as interactions with environmental factors and their respective stressors.

Primary versus Secondary Stress Pathways in RA

A fundamental objective in clinical research is to differentiate primary determinant factors from secondary reactions in the development and course of disease. Both the presymptomatic and clinically active phases of disease need to be critically investigated. Additionally, a priority challenge is to determine normal relations and interactions among the relevant physiological systems in health that become altered in the disease process. Generally, little research data are available on such early physiological interactions in clinically healthy subjects or in susceptible people before clinical onset. Most available data pertain to the pathological alterations that occur during active phases of clinical disease, without knowing the physiopathogenetic pathways (Figure 3).

Stressors may Amplify Dysregulations in RA

Under normal circumstances, counterregulatory balance occurs between the NEI systems. Once disease manifests, definite immunological activation and inflammation become evident, which, in turn, stimulate the HPA axis. In active disease, blood levels of the main inflammatory cytokines are significantly elevated (i.e., interleukin (IL)-1β, IL-6, and tumor necrosis factor (TNF)-α), and they centrally stimulate the HPA axis. The HPA axis products normally act to counterregulate overactivation of inflammatory pathways (Figure 3). However, such HPA axis actions are suspected to be relatively deficient in some people susceptible to RA, particularly younger women with onsets in their premenopausal ages.

In humans, cortisol is secreted from the adrenal cortex as the main anti-inflammatory product of the HPA axis, whereas in rodents, corticosterone is the main glucocorticoid. Blood levels of these corticosteroids are tightly controlled by negative feedback

mechanisms at the hypothalamic and pituitary (HP) levels. The concept has evolved that relative deficiency or insufficiency of glucocorticoid responsiveness permits inappropriately exaggerated activation of the immunological and microvascular endothelial systems. Increased predisposition to inflammatory disorders may also result from natural susceptibilities via other host immunological mechanisms or added predispositions from varied stressors.

Firm data are not available on mechanisms or pathways whereby particular stressors can alter the normal psychoneuroendocrine physiology and influence susceptibly to RA. However, one study of 2490 Vietnam Theater Veterans diagnosed as having PTSD showed a fivefold increased likelihood to have postwar RA. That study and other data may indicate that people with relatively lower HPA axis and SNS stress responsiveness may be more susceptible to both PTSD and autoimmune disorders, like RA. Alternatively, stressors that predispose to PTSD may suppress HPA axis responsiveness via inhibition of corticotropin-releasing hormone (CRH). In turn, counterregulation of immunological activation could be diminished. Research on presymptomatic risk factors for onset of RA indicate that increased predisposition is related to evidences of: (1) increased immunological activation, (2) physiologically lower adrenal cortical reserve in younger women, and (3) history of heavy cigarette smoking in females and males, all of which may impact stress responsiveness of the NEI and MVE systems.

An Evolving Research Challenge in Rheumatoid Arthritis

In growth, maturation, and later aging into the senescent period, marked changes normally occur in the NEI and MVE systems. Attempts have been made to correlate such important physiological changes with risks of developing RA in the population. In women, incident risks of RA onset increase rather regularly with aging, from adolescence to the people in their seventies. However, in males aged 15–50 years, the risks are level and low, about one-fifth as high as in females. Such differential susceptibility patterns indicate that males are relatively protected during their younger adult and more masculine years. From such data, adrenal and gonadal androgens have been implicated as being protective of RA, perhaps by exerting greater control of immunological activation.

Estrogens control endothelial cell activation more effectively than androgens. Premenopausal women have significantly lower natural risk of coronary artery disease (CAD) and myocardial infarction (MI) than male counterparts. Sufficient intensity and duration of cigarette smoking exposures are stressful vasculotoxins that increase risks of CAD and RA in both sexes. The smoking effect on CAD is magnified in women by use of oral contraceptives (OCs), which provide supraphysiological or pharmacological doses of estrogens. However, OCs alone have not been proven to amplify risks of RA. Also, pregnancy typically improves RA activity, which argues against female sex hormones being a primary risk factor.

Pharmacological doses of estrogens have been shown to increase risks of CAD and MI in both females and males. However, the mechanism is likely to be due to increased thrombogenic risks, rather than by microvascular endothelial or inflammatory pathways. Interpretation of such clinical-epidemiological risk patterns are difficult. Nevertheless, tentative hypotheses may be drawn.

First, heavy cigarette smoking is a stressor which increases risks for a large spectrum of microvascular-related disorders, including RA, and equally so in males and females.

Second, males have significantly lower natural risk of RA during their younger adult ages, suggesting greater control of immunoactivation during those years of highest androgenicity. A significant association of PTSD and RA is reported in Vietnam Theater Veterans, suggesting that stress or diminished tolerance thereto may be a susceptibility factor. Typically, stress reactions result in decreased androgen levels, which could affect susceptibility to RA in males.

Third, estrogen levels *per se* are not correlated with increased RA risks, because younger and pregnant women have considerably lower incidence than postmenopausal women. Also, OCs have not been proven to increase RA risk. Thrombogenic pathways, enhanced by pharmacological doses of estrogens, do not seem to be a primary mechanism in RA, in contrast to immunoactivation and MVE perturbations.

Summary

Interfaces of the immunological and microvascular endothelial systems seem to be primarily incriminated in the inflammatory pathways in RA. Their activation may be amplified permissively by relatively insufficient HPA axis counterregulatory controls. Sufficient cortisol responsiveness is needed to control capillary permeability and transmigration of inflammatory cells into the interstitium. This potent hormone is also required to control overactivation of the immunological system and inflammatory cytokine production, which are key mediators in the RA process.

Stressors and environmental activators may be suspected as being contributory co-factors in the complex web of causation of RA (Figure 3). Heavy cigarette smoking has known microvascular toxic effects. Also, stressful life circumstances likely compromise

normal HPA axis counterregulatory control functions and diminish androgen levels, particularly in males. The glucocorticoids, cortisol in humans and corticosterone in rodents, are the most potent anti-inflammatory products of the HPA system and in the stress response. However, many additional steroid hormones are secreted by the adrenal cortex, in addition to those products of the gonads. Also, catecholamines are secreted by the adrenal medulla and ANS, which are normally immunosuppressive. Besides glucocorticoids, these additional products of the HPA axis, gonads, and ANS can also modulate activation of the immunological–endothelial systems and control their interactions. Further research is needed on the complex relations of stress reactions to risks and course of RA and other rheumatic disorders.

Acknowledgments

Support for this project was provided by the Department of Medicine, University of Illinois College of Medicine at Peoria, and a grant from the MTM Foundation.

Further Reading

Boscarino, J. A. (2004). Posttraumatic stress disorder and physical illness. *Annals of the New York Academy of Sciences* **1032**, 141–153.

Cannon, W. B. (1929). Organization for physiological homeostasis. *Physiological Reviews* **9**, 399–431.

Felson, D. T. and Zhang, Y. (1998). An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. *Arthritis and Rheumatism* **41**, 1343–1355.

Freeman, W. J. (2003). Neurodynamic models of brain in psychiatry. *Neuropsychopharmacology* **28**, S54–S63.

Masi, A. T. (2000). Hormonal and immunologic risk factors for the development of rheumatoid arthritis: an integrative physiopathogenetic perspective. *Rheumatic Disease Clinics of North America* **2**, 775–803.

Masi, A. T. and Aldag, J. C. (2005). Integrated neuroendocrine immune risk factors in relation to rheumatoid arthritis: should rheumatologists now adopt a model of a multiyear, presymtomatic phase? *Scandinavian Journal of Rheumatology* **34**, 342–352.

Masi, A. T. and Chang, H. J. (1999). Cigarette smoking and other acquired risk factors for rheumatoid arthritis. In: Kaufman, L. D. & Varga, J. (eds.) *Rheumatic diseases and the environment*, pp. 111–127. New York: Chapman & Hall.

Masi, A. T. and Walsh, E. G. (2003). Ankylosing spondylitis: integrated clinical and physiological perspectives. *Clinical and Experimental Rheumatology* **21**, 1–8.

Masi, A. T., Aldag, J. C. and Jacobs, J. W. G. (2005). Rheumatoid arthritis: neuroendocrine immune integrated physiopathogenetic perspectives and therapy. *Rheumatic Disease Clinics of North America* **31**, 131–160.

Masi, A. T., DaSilva, J. A. P. and Cutolo, M. (1996). Perturbations of hypothalamic-pituitary-gonadal (HPG) axis and adrenal androgen (AA) functions in rheumatoid arthritis. *Baillière's Clinical Rheumatology* **10**, 295–332.

Masi, A. T., White, K. P. and Pilcher, J. J. (2002). Person-centered approach to care, teaching, and research in fibromyalgia syndrome: justification from biopsychosocial perspectives in populations. *Seminars in Arthritis and Rheumatism* **32**, 71–93.

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* **338**, 171–179.

Robin, E. D. (1979). Claude Bernard. Pioneer of regulatory biology. *JAMA* **242**, 1283–1284.

Selye, H. (1936). Thymus and adrenals in the response of the organism to injuries and intoxications. *British Journal of Experimental Pathology* **17**, 234–240.

Selye, H. (1946). The general adaptation syndrome and the diseases of adaption. *Journal of Clinical Endocrinology and Metabolism* **6**, 117–230.

Straub, R. H. and Besedovsky, H. O. (2003). Integrated evolutionary, immunological, and neuroendocrine framework for the pathogenesis of chronic disabling inflammatory diseases. *FASEB Journal* **17**, 2176–2183.

Yunus, M. B. (2005). The concept of central sensitivity syndromes. In: Wallace, D. J. & Clauw, D. J. (eds.) *Fibromyalgia and other central pain syndromes*, pp. 29–44. Philadelphia: Lippincott Williams & Wilkins.

Walker, J. G., Littlejohn, G. O. and McMurray, N. E., et al (1999). Stress system response and rheumatoid arthritis: a multilevel approach. *Rheumatology* **38**, 1050–1057.

Ribosomes *See:* Protein Synthesis; Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Response, Overview; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Risk Factors for Stress *See:* Life Events Scale; Life Events and Health.

S

Salivary Cortisol

C Kirschbaum and D H Hellhammer
Technical University of Dresden, Dresden, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 379–383, © 2000, Elsevier Inc.

Methodological Issues
Salivary Cortisol Stress Responses
Field Studies and Ambulatory Assessments
Psychoendocrine Stress Research in Infants and Children

The hypothalamic-pituitary-adrenal (HPA) axis responds rapidly and rather specifically to a wide range of environmental and internal demands often referred to as stress. It is thought that the HPA response to stress plays a pivotal role in the organism's attempt to maintain function through change, as expressed in the allostasis model. Although the HPA hormones corticotropin releasing hormone (CRH) and adrenocorticotropic hormone (ACTH) may exert additional adaptive effects, accumulating evidence suggests that cortisol is the main hormone responsible for allostatic stress responses.

Following stimulation by ACTH, cortisol is synthesized and secreted from the adrenal glands and released into the circulating blood, where it is bound rapidly to carriers such as cortisteroid-binding globulin (CBG), albumin, and erythrocytes. Only a small fraction (2–15%) of cortisol released remains unbound, or free. According to the free hormone concept, it is only this hormone fraction that brings about the multitude of genomic cortisol effects in peripheral tissues and in the brain. Although in blood both bound and free cortisol can be measured, only free cortisol appears in saliva. Hence, the measurement of salivary cortisol provides an index of the biologically active fraction of this steroid hormone.

Methodological Issues

For decades, research on the acute and chronic effects of stress has employed cortisol levels as an index of the individual response to stress. Due to methodological reasons, this research was slowed by the need for multiple blood samples to measure cortisol levels repeatedly. Blood sampling is not only labor-intensive (and expensive) but sometimes also counterproductive in stress research because the procedure of venipuncture itself can lead to significant HPA activation. The advent of advanced biochemical assays capable of measuring cortisol reliably in the lower nanomolar range, however, helped to overcome this problem.

Since the mid-1980s, the assessment of free cortisol in saliva has become an increasingly important tool in stress research because it has a number of advantages over the measurement of cortisol in blood or urine. Of prime interest to the stress researcher is the noninvasiveness of sampling and the ability to obtain samples at short intervals from subjects over the entire life span without raising ethical problems associated with venipuncture. The measurement of cortisol levels in saliva is the method of choice for field studies or ambulatory assessments in the natural environment of the subjects because medical personnel or laboratory facilities are not required for collecting and storing the samples. Salivary cortisol is not measured only in humans, it also serves as an endocrine correlate of stress in animals, including pigs, monkeys, elephants, sheep, tree shrews, guinea pigs, and goats.

Sampling, Storage, and Biochemical Analysis

Samples are usually obtained by the subjects or patients; infants and young children (<2 years old) require additional help from parents or hospital staff. Subjects either salivate directly into small containers or use devices such as the Salivette, which is currently the most frequently employed sampling method.

Stimulation of saliva flow by gently chewing on the device or by adding a few crystals of sugar-sweetened Kool-Aid (for use with infants and small children) ensures adequate sample volume. Because most immunoassays available today provide false high values with low pH samples, saliva should not be sampled after stimulation with citric acid. The storage and mailing of samples are uncomplicated. Because cortisol is rather stable in saliva, samples can be stored at room temperature for at least 4 weeks without a significant reduction in cortisol concentration or mailed to a laboratory at ambient temperature, even overseas. However, in order to keep samples over prolonged periods and to avoid molding, it is recommended that samples be stored at −20°C or lower. For the biochemical analysis of salivary cortisol, several methods are now available that allow reliable measurement of the steroid hormone. Immunoassays with radioactive and other tracers are the method of choice for most laboratories. An increasing number of commercial assay kits designed for use with serum or plasma now also provide instructions for use with saliva. In a few instances, the use of alternative analytical techniques may be required. For example, the simultaneous detection of several steroid hormones and their metabolites in saliva can be achieved with high-pressure liquid chromatography (HPLC). Also, for processing samples that contain high levels of compounds (e.g., prednisone) that cross-react significantly with most antibodies raised against cortisol for use with immunoassays, HPLC may be the adequate analytical technology. Gas chromatography-mass spectrometry is used only for validating cortisol results obtained with other assay methods.

Comparison with Cortisol Levels in Blood

Because cortisol is a small (molecular weight, MW, 362) and highly lipid-soluble molecule, the unbound hormone can pass easily through the lipid-bilayer membranes of nucleated cells. This allows free cortisol to appear in all bodily fluids, including blood, cerebral spinal fluid, urine, sweat, semen, and saliva. Cortisol bound to carriers is usually excluded from these body compartments.

Most important for stress research, cortisol levels measured in saliva agree very well with the amount of free cortisol in blood. Correlations between salivary and unbound blood cortisol levels usually explain more than 80% of the total variance observed ($r \geq 0.90$). This high agreement is due to the fact that cortisol enters the cell and the oral cavity mainly by passive diffusion. It is therefore independent of transport mechanisms and saliva flow rate, in contrast to other components also found in saliva (e.g., immunoglobulin A and catecholamines). Absolute levels of free cortisol are lower in saliva due to a relative abundance of the cortisol-metabolizing enzyme 11-β hydroxysteroid dehydrogenase.

Circadian Rhythm and Normal Values

As in blood, salivary cortisol levels follow a robust circadian rhythm that has been studied only in awake subjects so far. Peak levels are usually found in the early morning hours, with decreasing values thereafter. Several secretory episodes can be detected throughout the day, even in the absence of psychological or physical stress.

Stimulated by an individual's waking up in the morning, there is a 50–100% increase in cortisol levels, peaking about 30 min after awakening. This response shows remarkable stability over weeks and months in the same individual and has been found to relate to prolonged stress exposure and the feeling of burnout. After the morning peak, cortisol levels decrease steadily until a second major secretory episode occurs in response to larger meals at lunch time. As with the wakening response, meal-related peaks appear to be more pronounced (approximately 1.5-fold) in free cortisol compared to total blood cortisol. During afternoon and evening hours, saliva cortisol continues to decrease with no further major secretory episode. Figure 1 depicts the circadian rhythm of cortisol in saliva representing normal values typically observed in a healthy population.

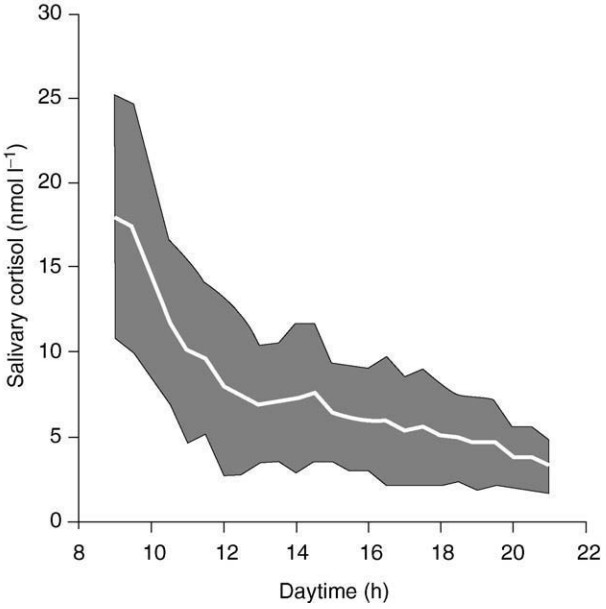


Figure 1 Normal values of salivary cortisol during the day (white line: mean levels; shaded area: means ±1 SD).

Salivary Cortisol Stress Responses

Response Kinetic

In contrast to other neuronal and endocrine response circuits, it takes several minutes before an HPA stress response becomes visible in the periphery. After the onset of a psychosocial stressor or strenuous exercise, ACTH levels start to rise within less than 5 min and show peak values shortly after the cessation of stress (<5 min). The cortisol response lags by 5–20 min, with maximal cortisol levels 10–30 min after stress cessation. Although cortisol exerts a strong feedback action at the level of the pituitary, hypothalamus, and hippocampus, prolonged periods of stress can lead to cortisol secretion over several hours. For example, in marathon runners, salivary cortisol was observed to increase monophasically for more than 4 h, peaking approximately 30 min after the run was finished.

The transfer of cortisol from plasma to saliva occurs quickly. Within less than a minute, cortisol injected intravenously appears in saliva, and the peak concentration in saliva lags by less than 2–3 min compared to levels measured in blood. Thus, it is not surprising to find similar stress response kinetics in saliva and in blood. For example, total plasma cortisol and salivary cortisol levels rise in parallel in response to a laboratory stress task that consists mainly of an anticipation period (3 min), public speaking (5 min), and mental arithmetic (5 min) in front of an audience. Following the cessation of stress, however, free cortisol tends to increase further for another 15–20 min, whereas total cortisol levels are unchanged or start to decline (Figure 2). A similar response pattern was observed in plasma after meals, bicycle ergometry, and injection of ACTH(1–24) and during sleep.

Discrepancy between Salivary Cortisol and Total Plasma Stress Responses

Although salivary cortisol levels agree well with plasma-free cortisol, they show only moderate correlations with total cortisol levels under certain circumstances. Although the adrenal cortex was stimulated by ACTH and secreted significant amounts of cortisol, subjects may show blunted or completely absent cortisol responses as measured in saliva. This phenomenon can be observed in women using oral contraceptives (OC) (Figure 3). The use of oral contraceptives leads to an increase in CBG levels and hence to a larger number of binding sites for cortisol in the circulating blood. Therefore, more cortisol released from the adrenals is bound by CBG in these women than in non-OC users, and lower levels of the biologically active free fraction appear in saliva (as well as in target tissues). Similar discrepancies between total blood and salivary cortisol levels should be expected in all cases of increased or decreased CBG production.

Field Studies and Ambulatory Assessments

One of the major advantages of salivary cortisol is that samples can be obtained in the field, that is, in the natural (social) environment of the subjects/patients

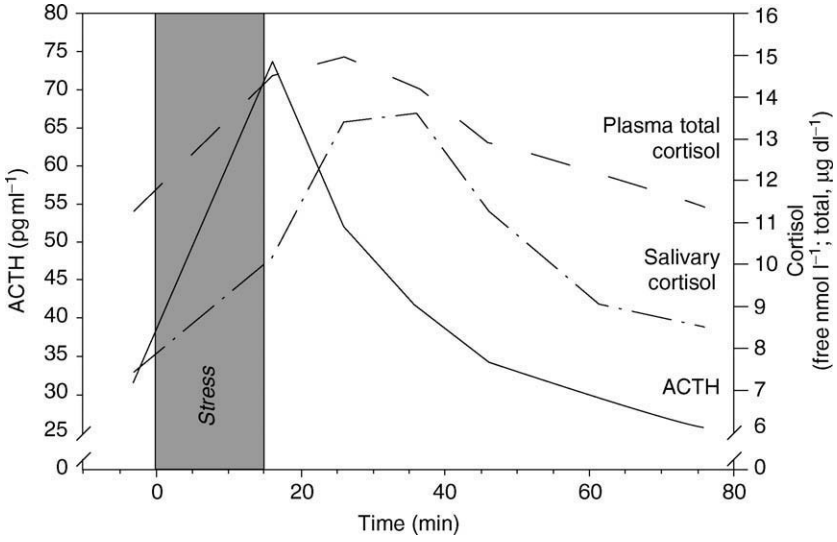


Figure 2 Mean ACTH, total plasma cortisol, and salivary cortisol responses to a brief psychosocial stress test (Trier social stress test). Note that there are different units of hormone concentrations for total plasma and salivary cortisol.

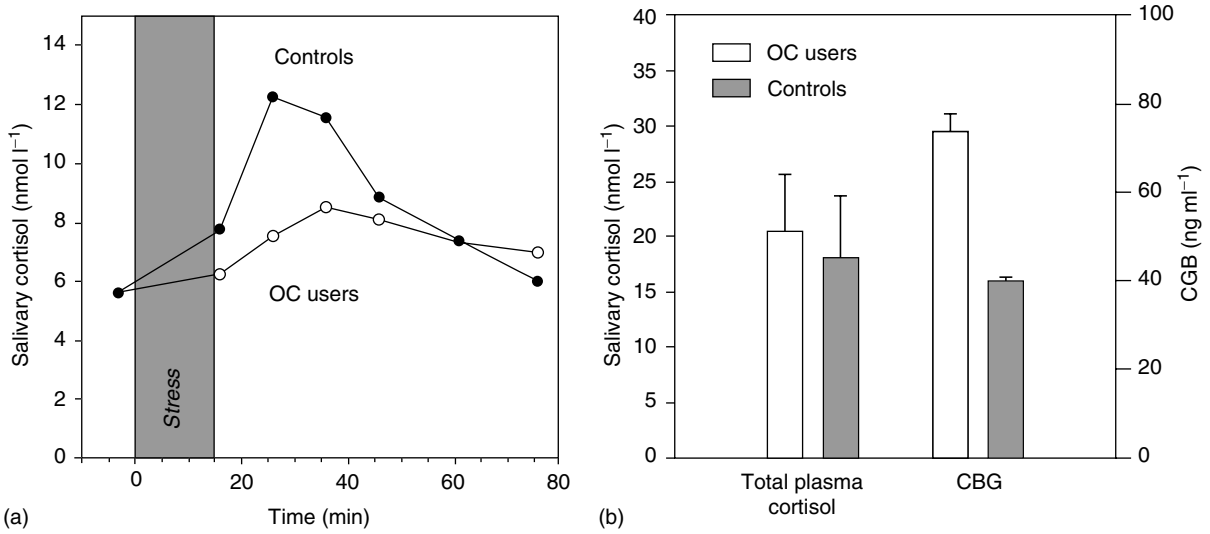


Figure 3 Cortisol responses to stress in women using oral contraceptives (OC) and nonusers. a, Salivary cortisol levels; b, Total plasma cortisol and corticosteroid-binding globulin (CGB) levels. The blunted free cortisol response pattern in OC users is not due to lower cortisol secretion but rather a consequence of enhanced CGB levels. AUC, area under the curve.

as well as in special settings outside a psychological or medical laboratory. This advantage has allowed researchers to perform a number of studies that would have been very difficult or even impossible with blood sampling. With saliva, large cross-sectional studies can be performed by mailing sampling devices and instructions to the subjects, who return the material by regular mail to the laboratory. Populations of 500–1000 subjects or more can be investigated easily by this method.

Of special interest to the stress researcher is the ease of obtaining cortisol readings from subjects repeatedly before and after exposure to real-life stress situations. For example, the impact of the first three parachute jumps on acute HPA responses and subsequent pharmacological tests was studied with saliva samples obtained at the airfield at 20-min intervals over 10 h. In addition, the same subjects collected saliva samples at identical times on a control day to relate the stress-induced changes to the individuals' normal circadian rhythms. Furthermore, ambulatory (or momentary) assessments of cortisol responses to daily stress can be performed independently of the researcher. Using electronic devices such as preprogrammed wristwatches and pocket-size (computer) booklets, researchers can beep study subjects several times a day and request a report on current, past, or anticipated stressors. Twenty minutes later, they beep the subjects again and ask them to take a saliva sample. Several studies with momentary assessments of stress and salivary cortisol have been published supporting the view that even minor changes in negative affect can result in increased cortisol levels, whereas a positive affect has the opposite effect.

Psychoendocrine Stress Research in Infants and Children

For ethical reasons, invasive measures are usually prohibited for use with healthy infants or children. Circumventing these difficulties, the availability of salivary cortisol has stimulated research into the psychoendocrinology of stress in infants and children significantly. A rapidly growing literature shows that the human HPA axis begins to respond to psychosocial or physical stimulation with an activation of the axis very early in life. In combination with appropriately modified protocols, the acute response to experimental stress can thus be monitored in children and compared directly with adult response patterns, which will eventually lead to a more detailed understanding of the ontogenesis of the human HPA stress response.

Acknowledgments

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (Ki 537/6-1).

See Also the Following Articles

Allostasis and Allostatic Load; Hypothalamic-Pituitary-Adrenal; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress.

Further Reading

Brooks, F. S. and Brooks, R. V. (1984). Cortisol and cortisone in saliva. In: Read, G. F., Riad-Fahmy, D., Walker, R. F. & Griffiths, K. (eds.) *Immunoassays of steroids in saliva*, pp. 322–326. Cardiff, UK: Alpha Omega.

Buske-Kirschbaum, A., Jobst, S., Wustmans, A., et al. (1997). Attenuated cortisol response to psychosocial stress in children with atopic dermatitis. *Psychosomatic Medicine* **59**, 419–426.

Cook, N. J., Read, G. F., Walker, R. F., Harris, B. and Riad-Fahmy, D. (1992). Salivary cortisol and testosterone as markers of stress in normal subjects in abnormal situations. In: Kirschbaum, C., Read, G. F. & Hellhammer, D. H. (eds.) *Assessment of hormones and drugs in saliva in bio-behavioral research*, pp. 147–162. Seattle, WA: Hogrefe & Huber.

Deinzer, R., Kirschbaum, C., Gresele, C., et al. (1997). Adrenocortical responses to repeated parachute jumping and subsequent h-CRH challenge in inexperienced healthy subjects. *Physiology & Behavior* **61**, 507–511.

Ekins, R. (1990). Measurement of free hormones in blood. *Endocrine Review* **11**, 5–46.

Follenius, M. and Brandenberger, G. (1986). Plasma free cortisol during secretory episodes. *Journal of Clinical Endocrinology and Metabolism* **62**, 609–612.

Gunnar, M. R. (1989). Studies of the human infant's adrenocortical response to potentially stressful events. *New Directions in Child Development*, 3–18.

Kirschbaum, C. and Hellhammer, D. H. (1989). Salivary cortisol in psychobiological research: an overview. *Neuropsychobiology* **22**, 150–169.

Kirschbaum, C., Kudielka, B. M., Gaab, J., et al. (1999). The impact of gender, menstrual cycle phase and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosomatic Medicine* **67**, 154–162.

Kirschbaum, C., Pirke, K.-M. and Hellhammer, D. H. (1993). The "Trier Social Stress Test" – a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* **28**, 76–81.

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* **338**, 171–179.

Pruessner, J., Wolf, O. T., Hellhammer, D. H., et al. (1997). Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. *Life Sciences* **61**, 2539–2549.

Spangler, G. and Scheubeck, R. (1993). Behavioral organization in newborns and its relation to adrenocortical and cardiac activity. *Child Development* **64**, 622–633.

van Eck, M., Berkhof, H., Nicolson, N., et al. (1996). The effects of perceived stress, traits, mood states, and stressful daily events on salivary cortisol. *Psychosomatic Medicine* **58**, 447–458.

Walker, R. F., Joyce, B. G., Dyas, J. and Riad-Fahmy, D. (1984). Salivary cortisol. I: Monitoring changes in normal adrenal activity. In: Read, G. F., Riad-Fahmy, D., Walker, R. F. & Griffiths, K. (eds.) *Immunoassays of steroids in saliva*, pp. 308–316. Cardiff, UK: Alpha Omega.

Salt Appetite

R S Weisinger
La Trobe University, Victoria, Australia
N Chen
University of Melbourne, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R S Weisinger and D A Denton, volume 3, pp 384–392, © 2000, Elsevier Inc.

Introduction
Environmental Stress and Salt (Sodium Chloride) Appetite
Hormones of the Hypothalamic-Pituitary-Adrenal Axis Associated with Stress and Salt Appetite
Brain Mechanisms Involved in Stress-Induced Salt Appetite
Conclusion

Glossary

Homeostasis	The maintenance of a relatively constant internal environment.
Hypertension	Abnormally or excessively high blood (arterial) pressure.
Hypothalamic-pituitary-adrenal (HPA) axis	The integrated association between the hypothalamus (an area of the brain that integrates the various automatic body reactions), the pituitary gland (a small gland at the base of the brain that controls hormone release from the adrenal gland), and the adrenal gland.
Ingestive behavior	Taking food or fluid into the body as nourishment.
Lesion	The removal of or irrevocable damage to a localized region of body tissue.

Sympathetic nervous system One of the two divisions of the autonomic nervous system. It innervates involuntary processes, which are largely beneath consciousness (e.g., the actions of the visceral organs, heart, blood vessels, glands, and smooth muscle). The activation of the sympathetic nervous system prepares the body for fight or flight and works antagonistically with the second autonomic division, the parasympathetic nervous system.

Introduction

Stress is often seen as a major contributor to disease, with specific effects on several of the body's physiological and behavioral systems. Its effect on the cardiovascular system may result in increased mortality. In addition, stress or the hormonal consequences of stress cause several changes in ingestive behavior. Evidence suggests that stress or the hormones of stress can directly influence the intake of salt (sodium chloride). An increased intake of salt is widely accepted to be a factor in the causation of hypertension and, in some instances, it magnifies the cardiovascular effects of stress. Evidence regarding the brain mechanisms involved in the changes in salt appetite caused by stress and the hormones of stress are described.

Environmental Stress and Salt (Sodium Chloride) Appetite

Environmental stress can result in increased salt appetite, as has been shown in several ways. One of the first demonstrations was made while attempting to fit jackets, designed for carrying small infusion pumps, on wild rabbits (see Figure 1). This minor restraint on free movement caused the rabbits to become agitated, and, over the course of several days, a spectacular increase in salt appetite was observed – their intake of a concentrated salt solution increased by over 15-fold. This intake of salt approximated a daily turnover of 30% of extracellular fluid sodium, sodium in the tissue fluid outside the cells. In similar experiments in which restraints were placed on mice (see Figure 2), an increase in salt intake was also observed. This salt intake was eliminated by the subcutaneous administration of the drug captopril; this drug prevents the formation of the peptide hormone angiotensin II. Angiotensin II, produced peripherally (i.e., in the circulation) and in the brain, is involved in both sodium and cardiovascular homeostasis, and it is one of the hormones known to be elevated in stressful situations. Increased levels of angiotensin II are also thought to explain the increased salt intake of mice kept in a cold environment for several hours a day. It is well known that rats from one colony become over-reactive when introduced into another colony. An increased salt appetite has been documented in these stressed animals.

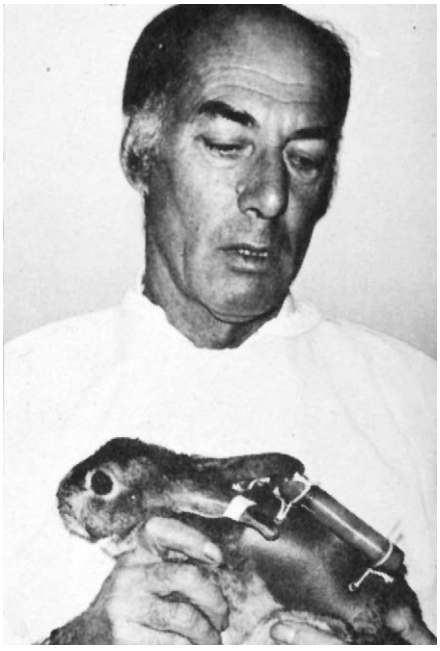


Figure 1 Jack Nelson holding a wild rabbit in a jacket designed to carry infusion pumps.

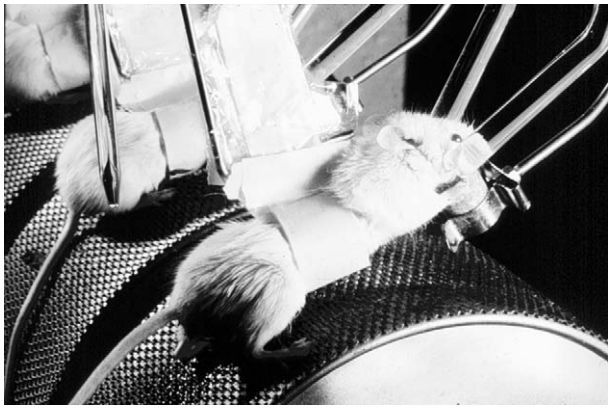


Figure 2 Restrained mouse on running wheel.

The sympathetic nervous system is responsible for the fight or flight response in animals confronted with stressful situations. Drugs such as clonidine or reserpine, which interfere with the actions of the sympathetic nervous system, block the characteristic stress-induced increase in salt intake. This suggests that the increase in salt intake involves the increased activity of the sympathetic nervous system.

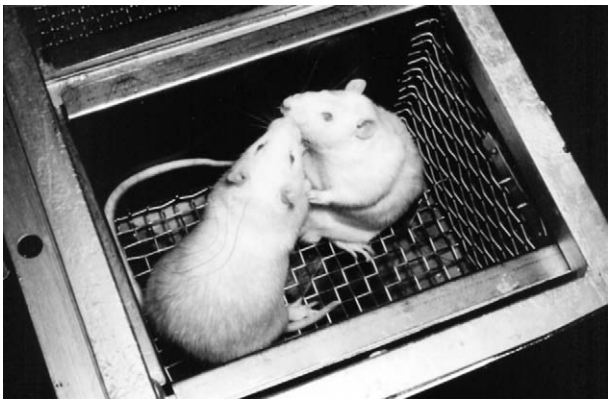


Figure 3 Two adult male Sprague-Dawley rats in conflict during crowding.

Crowding, an imposed social stress, was also found to increase salt intake in rats. Following a control period during which rats were housed alone in a small individual cage, two rats were placed together into a different cage (see Figure 3). The intake of concentrated salt solution increased over several weeks.

In summary, animals maintained in stressful situations have been shown to manifest an increase in salt intake. This increased salt intake is mediated by increased levels of the hormones associated with stress, namely, the hormones of the hypothalamic-pituitary-adrenal (HPA) axis, primarily adrenocorticotropic hormone (ACTH) and the mineralocorticoid/glucocorticoid hormones secreted by the adrenal gland. Angiotensin II, another hormone whose levels are increased in stressful situations, can stimulate the secretion of ACTH and the adrenal hormones. Interfering with the actions of angiotensin II can block salt appetite induced by stress. At present, the contribution of the sympathetic nervous system to salt appetite induced by stress seems paradoxical given that the activation of this system causes increased blood pressure.

Hormones of the Hypothalamic-Pituitary-Adrenal Axis Associated with Stress and Salt Appetite

It is well documented that hormones of the HPA axis (see Figure 4) can cause salt appetite. These include corticotropin releasing hormone (CRH) from the paraventricular nucleus of the hypothalamus, ACTH from the pituitary gland at the base of the brain, and the adrenocortical hormones (often referred to as steroid hormones), including the mineralocorticoid hormones, aldosterone and deoxycorticosterone and the glucocorticoid hormones, corticosterone and cortisol. The name of the former suggests their important role in the retention of minerals (i.e., salt) in the body and the name of the latter reflects their role in glucose metabolism. CRH and the more recently discovered urocortin (UCN) are thought to have a major physiological role in the response to stress, including the stimulation of ACTH secretion. ACTH and angiotensin II are important in stimulating the release of the mineralocorticoid hormones from the adrenal cortex.

Adrenocortical Hormones

There are at least 20 known steroid hormones with numerous functions as well as catecholamines produced and secreted by the adrenal gland. Work, started over 50 years ago, has demonstrated that the administration of mineralocorticoid hormones (aldosterone or deoxycorticosterone) causes salt appetite. Mineralocorticoid-induced salt appetite has been shown in rats, mice, rabbits, hamsters, and sheep. Sheep treated with deoxycorticosterone shift their preference from a low-salt diet to a high-salt diet, and the increased intake of salty food is associated with an increase in both water intake and blood pressure. In addition, infusion of aldosterone into the brain causes increased blood pressure in rats.

Mineralocorticoid hormones may also influence salt appetite in the rat by its interaction with angiotensin II. The synergy between the mineralocorticoid hormones and angiotensin II may be due to the ability of the mineralocorticoid hormones to increase the number or sensitivity of angiotensin II receptors. Evidence contrary to the proposal that salt appetite is caused by a synergy between mineralocorticoids and brain angiotensin II has been reported in sheep. Data regarding the synergy hypothesis are inconclusive in the pigeon and rabbit.

The administration of either cortisol or corticosterone increased salt appetite in rabbits and mice. The concurrent administration of both cortisol and corticosterone increased appetite by more than the sum of the individual intakes. This result suggests a synergy between the two glucocorticoid hormones. In humans, the administration of cortisol increased blood pressure but did not alter salt preference.

The glucocorticoid hormones may contribute to salt appetite by potentiating the appetite caused by mineralocorticoid treatment. The administration of corticosterone or dexamethasone, a potent synthetic glucocorticoid hormone, for example, enhances salt appetite induced by aldosterone or deoxycorticosterone in rats. A similar influence has been reported in mice.

In experiments in sheep, the intravenous infusion of a combination of steroids, a treatment that mimics

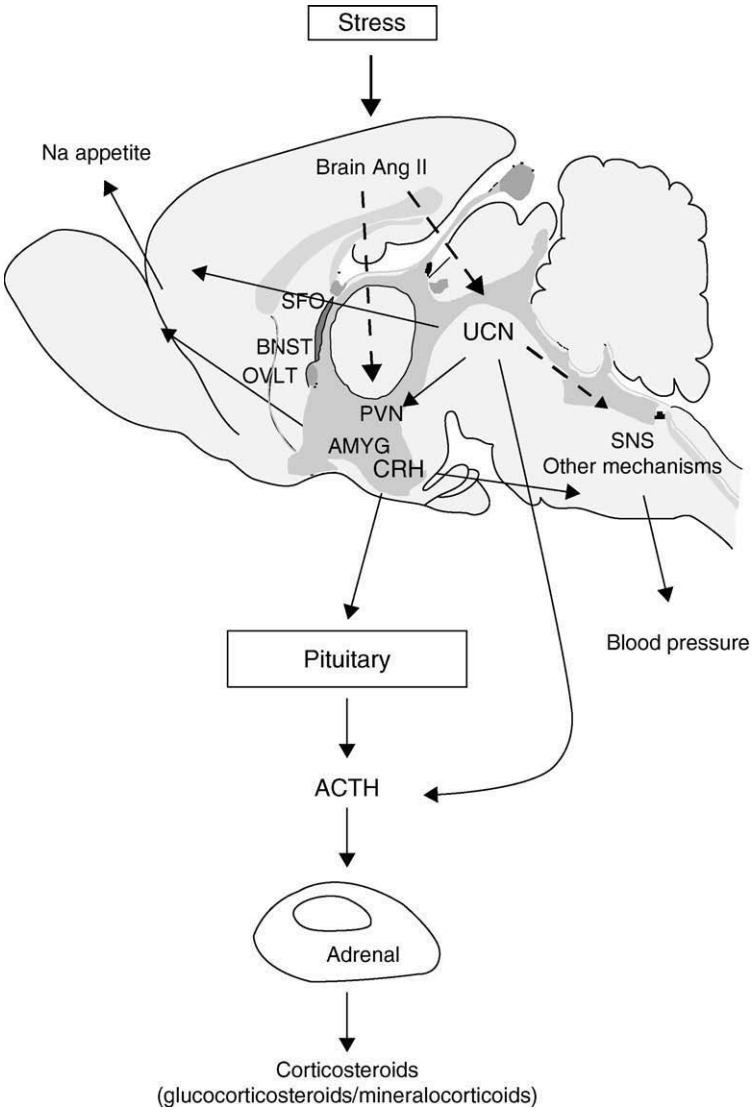


Figure 4 Some of the mechanisms affecting salt appetite and blood pressure. ACTH, adrenocorticotropic hormone, AMYG, amygdala (a laterally situated structure shown in central perspective), Ang II, angiotensin II; BNST, bed nucleus of the stria terminalis; CRH, corticotropin releasing hormone; OVLT, organum vasculosum of the lamina terminalis; PVN, paraventricular nucleus of the hypothalamus; SFO, subfornical organ; SNS, sympathetic nervous system; UCN, urocortin.

the effects of ACTH administration, increased salt appetite by an amount similar to that observed in sheep treated with ACTH. Blood pressure was also increased. The steroids infused included both mineralocorticoid (aldosterone and deoxycorticosterone) and glucocorticoid (corticosterone, cortisol, and 11-deoxycortisol) hormones and produced blood levels of these hormones similar to those observed during ACTH administration. Blood pressure was also elevated in this instance. However, the infusion of the steroids into the arterial blood directly supplying the brain or directly into the brain did not increase blood pressure. This suggests that, in contrast to the influence of aldosterone on blood pressure in the rat, the steroid-induced increase in blood pressure in sheep is not due to a central effect of the steroids.

In summary, an increased intake of salt is produced by the administration of mineralocorticoid and/or glucocorticoid hormones. Several synergies between adrenal hormones in regard to the stimulation of salt appetite have been reported. Also, mineralocorticoid hormones may work in conjunction with angiotensin II to stimulate salt intake.

Adrenocorticotropic Hormone

ACTH was isolated from mammalian hypothalamus in the1940s–1960s; it has the following properties:

- It is a 39-amino-acid peptide, derived from pro-opiomelanocortin and found primarily in the pituitary gland but also found in the brain (e.g., the arcuate nucleus).
- Its levels are elevated during stress.
- It controls the adrenal corticosteroid responses to stress.

The administration of ACTH causes a large increase in salt intake in several species such as sheep, rabbits, rats, and mice. When rats were offered a choice among concentrated solutions of sodium chloride, potassium chloride, calcium chloride, and magnesium chloride, their appetite was primarily for the sodium chloride solution. The intake of water was also increased. By the fifth day of treatment, the salt intake (and output) was equivalent to the daily turnover of total body salt. In mice, the increase in salt intake caused by ACTH was not altered by treatment with captopril. The observation that the stress-induced (but not ACTH-induced) salt intake was blocked by captopril suggests that the stress-induced production of angiotensin II precedes the increase in ACTH; that is, in stressful situations, angiotensin II has an important role in the stimulation of ACTH secretion.

Angiotensin II, administered into the circulation or directly into the brain, can stimulate the secretion of ACTH, possibly by stimulating the secretion of CRH from the paraventricular nucleus of the hypothalamus. The increase in water intake caused by ACTH has been shown to occur in rats and mice that were offered only water to drink. Results showed that ACTH caused an increase in the extracellular fluid sodium concentration, which caused the enhanced water intake. The administration of ACTH into the circulation causes a large and sustained increase in blood pressure in sheep, rats, and humans. In humans, blood pressure was elevated even when the dietary salt intake was low; however, much larger increases in blood pressure occurred when salt intake was high. In both sheep and rat studies, ACTH-induced increases in salt appetite and blood pressure were eliminated by adrenalectomy, suggesting that this mechanism is dependent on the adrenal hormones. In sheep, ACTH treatment also increased the heart rate.

Corticotropin Releasing Hormone and Urocortin

CRH was isolated from sheep hypothalamus and characterized in the early 1980s; it has the following properties:

- It is a 41-amino-acid peptide, primarily from the paraventricular nucleus of the hypothalamus but also found in other brain nuclei and in peripheral tissues.
- Its levels are elevated during stress.
- It coordinates responses to stress: increases the synthesis and release of ACTH from the anterior pituitary and influences the autonomic nervous system (e.g., increases sympathetic activity).
- Its actions are mediated by two receptor subtypes: CRH-R1, CRH-R2 (G-coupled, seven-transmembrane, bound receptors).

UCN was isolated from rat midbrain and characterized in 1990s; it has the following properties:

- It is a 40-amino-acid peptide, found in the brain (e.g., in the Edinger–Westphal nucleus, lateral superior olive, lateral hypothalamus, and supraoptic nucleus) and peripheral tissues (e.g., the pituitary, heart, and gastrointestinal tract).
- Its levels are elevated during stress.
- It coordinates responses to stress.
- Its actions are mediated primarily by CRH-R2 receptors.

CRH and UCN are found in brain areas known to be important for body sodium homeostasis (e.g., the hypothalamus; septal area; organum vasculosum of the lamina terminalis, OVLT; and amygdala) as well as in systems associated with autonomic function (e.g., the parabrachial nucleus and locus coeruleus). CRH and UCN are increased in the brain by stress and are involved in the initiation of the behavioral and physiological responses to stress. The major actions of CRH and UCN include the release of ACTH and the activation of the sympathetic nervous system. CRH also causes increased locomotor activity and modulation of the immune response. UCN shares a 45% homology with CRH, has CRH receptor-binding characteristics, and has a potency suggesting that it may mediate some of the actions, particularly of a behavioral nature, attributed to CRH.

The actions of CRH and UCN appear to be mediated by at least three CRH receptors (-R1, -R2α, and -R2β) located in brain. CRH-R1 receptors are located in various regions of the cerebellum, several cortical areas, and the anterior and intermediate lobes of the pituitary. CRH-R2 receptors are located in the lateral septal nucleus, medial and cortical nuclei of the amygdala, bed nucleus of the stria terminalis, hippocampus, hypothalamic regions, and the serotonergic neurons of the midbrain raphe nuclei. CRH-R2 receptors are important in integrating hypothalamic neuroendocrine functions, as well as the autonomic and behavioral actions of central CRH. CRH-R2α receptors are almost exclusively in the brain; CRH-R2β receptors are found in brain and periphery. CRH-R2 receptors have a 6- to 40-fold greater affinity for UCN than for CRH.

The administration of ovine CRH in the brain of rabbits increased sodium intake, caused a sustained increase in both cortisol and corticosterone, and transiently increased blood pressure by 10 mmHg. The administration of CRH into the circulation did not influence salt intake. In recent experiments in sheep, however, prolonged central (see Figure 5) or peripheral (see Figure 6) administration of CRH or UCN failed to stimulate, and in fact inhibited, salt appetite despite stimulating the secretion of ACTH, as evidenced by the increased plasma levels of cortisol. Furthermore, the prolonged administration in the brain of CRH and UCN decreased food intake and need-free salt intake in sheep, rats, mice, and baboons. The decrease in salt intake in baboons was transient, but the decrease in food intake was sustained throughout the infusion period. The mechanisms responsible for the inhibition of salt appetite are unclear. In sheep, the prolonged central or peripheral administration of either CRH or UCN increased blood pressure, but only UCN increased heart rate.

The administration of CRH and UCN could be expected to increase salt intake due to their stimulatory influence on the secretion of ACTH or activity of the sympathetic nervous system. It appears that the hormones associated with stress may have both excitatory (ACTH and adrenal hormones) and

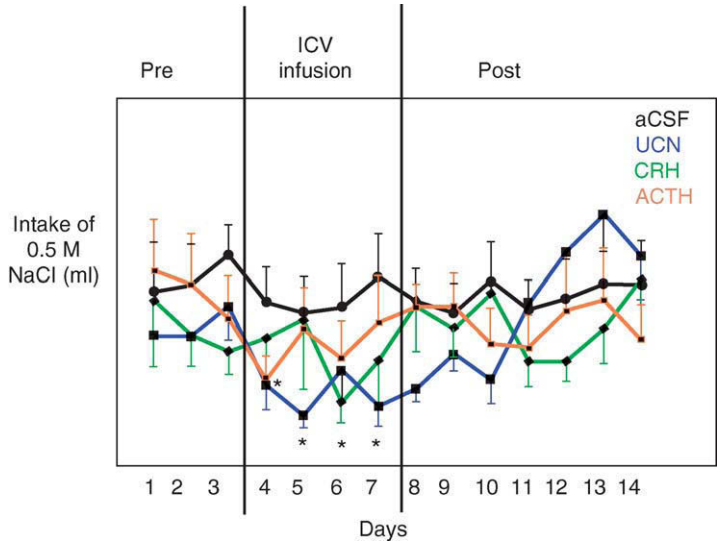


Figure 5 The effect of intracerebroventricular (ICV) infusions of aCSF, UCN, CRH, and ACTH on 0.5 M NaCl intake in sheep. * indicates $p < 0.05$. aCSF, artificial cerebrospinal fluid; ACTH, adrenocorticotropic hormone; CRH, corticotropin releasing hormone; UCN, urocortin.

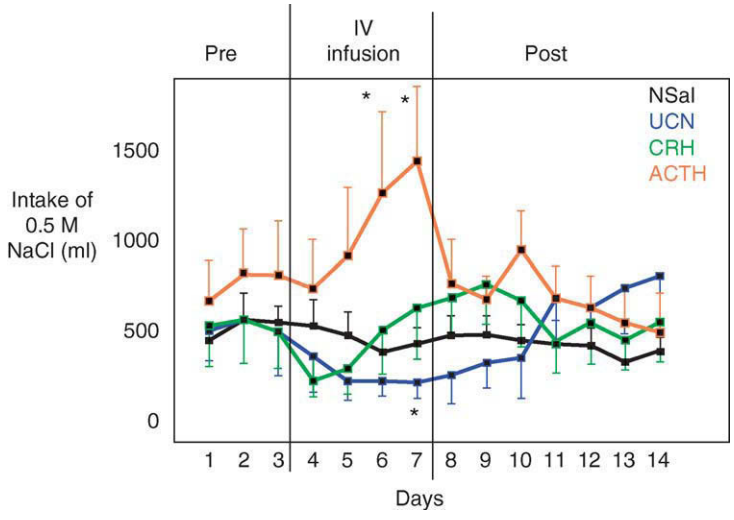


Figure 6 The effect of intravenous (IV) infusions of NSal, UCN, CRH, and ACTH on 0.5 M NaCl intake in sheep. ACTH, adrenocorticotropic hormone; CRH, corticotropin releasing hormone; NSal, normal saline; UCN, urocortin.

inhibitory (CRH and UCN) influences on salt intake. At present, the explanation for the inhibitory influence of CRH and UCN is not known. Presumably, the inhibition of salt appetite is caused by the influence of CRH on systems other than those influencing ACTH secretion or sympathetic nervous activity.

Combined infusion of Urocortin and Adrenocorticotropic Hormone

The observation that CRH and UCN increased the secretion of ACTH, as evidenced by the increase in plasma cortisol but decreased salt appetite, suggested that ACTH-induced salt appetite was inhibited. Administering the two peptides together tested the possibility that UCN inhibits ACTH-induced salt appetite. Indeed, the ACTH-stimulated salt intake was completely blocked by UCN, administered into either the brain or the circulation (see Figure 7). Additional observations indicated that UCN, given systemically, reduced the increase in blood pressure and enhanced the increase in heart rate caused by ACTH. In contrast, UCN, administered in the brain, did not alter the increase in blood pressure or the increase in heart rate caused by ACTH. The mechanisms responsible for the inhibitory role of UCN on salt appetite are still unknown. However, it is thought that UCN may play a role in steroid-induced salt appetite at the amygdala.

It is notable that prolonged peripheral treatment with a combined infusion of ACTH and UCN produced a marked increase in heart rate (see Figure 8), which was at least additive in magnitude when compared with heart-rate increases due to the infusion of either hormone alone. Plasma cortisol was also elevated with both the central and peripheral infusions of UCN in conjunction with the peripheral administration of ACTH. The major increase in heart rate in response to the combined administration of peripheral ACTH and UCN is probably via the direct action of CRH-R2 receptors on the heart. The inhibitory effect of UCN on blood pressure may be explained by the UCN-induced secretion of natriuretic peptides, molecules that increase excretion of sodium. These results indicate that sympathetic nervous system control of the cardiovascular system cannot be the only mechanism responsible for these stress hormone-induced changes. See Table 1.

Brain Mechanisms Involved in Stress-Induced Salt Appetite

Although the brain mechanisms responsible for salt appetite induced by stress have not yet been conclusively identified, there are several major candidates. The first involves midline brain tissue located on the front wall of the third ventricle (a donut-shaped, fluid-filled cavity). The circumventricular organs (i.e., the subfornical organ and the OVLT), notable for their lack of blood–brain barrier, are thought to be crucial to body fluid and mineral homeostasis in rats. A lesion to either of these areas causes a marked reduction in salt intake by the salt-depleted animal. However, sheep with lesions in this brain area responded normally to salt depletion. Evidence also indicates that these brain areas may contribute to salt appetite induced by treatment with the mineralocorticoid hormones. For example, there appear to be mineralocorticoid receptors within these brain regions and deoxycorticosterone treatment

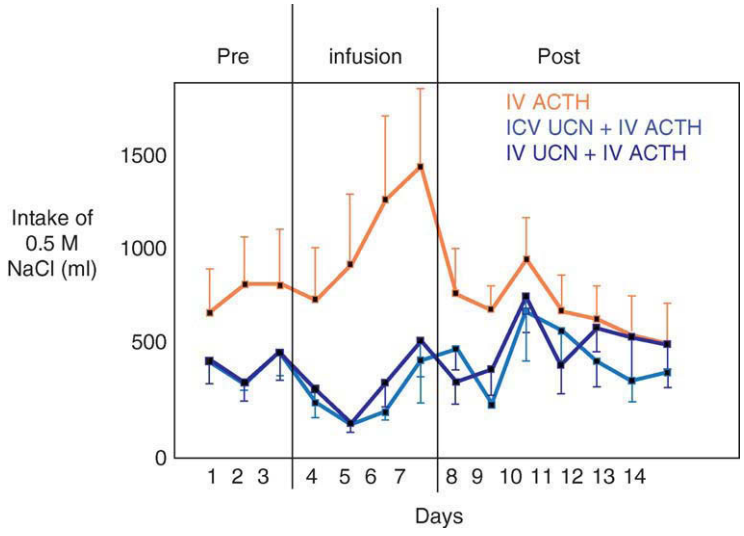


Figure 7 The effect of intracerebroventricular (ICV) and intravenous (IV) infusions of UCN on ACTH-induced 0.5 M NaCl intake in sheep. ACTH, adrenocorticotropic hormone; UCN, urocortin.

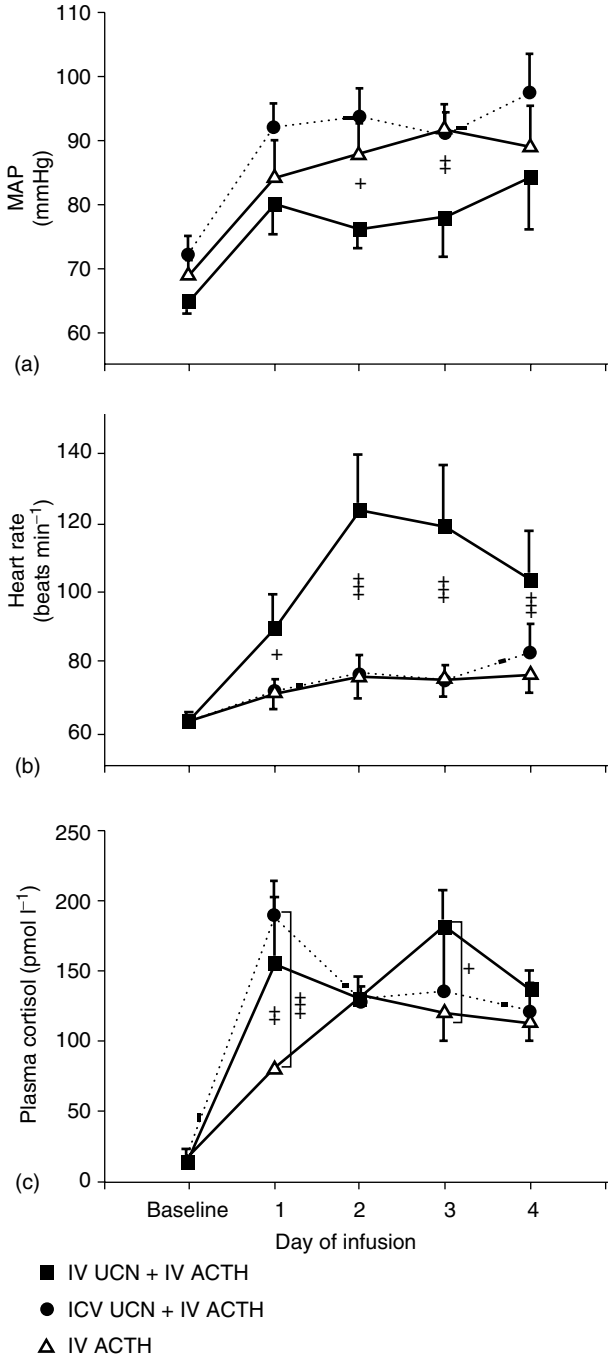


Figure 8 The effects of peripheral ACTH and central UCN infusions in sheep on a, blood pressure; b, heart rate; c, plasma cortisol concentration. ACTH, adrenocorticotropic hormone; ICV, intracerebroventricular; IV, intravenous; MAP, mean arterial pressure; UCN, urocortin.

increases neuronal activity at these brain areas. However, salt appetite induced by treatment with deoxycorticosterone is unchanged or enhanced in rats with lesions of this brain region.

The second candidate involves brain regions located some distance from the midline and behind the blood–brain barrier – the amygdala and bed nucleus of the stria terminalis. These brain areas are involved in the neural and hormonal responses to stress, especially those related to fear and anxiety. Increased neuronal activity has been reported in the medial amygdala of animals subjected to stress. Interference with mineralocorticoid receptor formation, with antisense to its mRNA, in the amygdala has been shown to interfere with the expression of deoxycorticosterone-induced salt appetite. Rats with lesion of the medial amygdala have impaired aldosterone and deoxycorticosterone-induced salt appetite, and the corticosterone potentiation of aldosterone-induced salt appetite is eliminated. Despite the evidence that salt appetite caused by salt depletion is mediated by the synergy between aldosterone and angiotensin II, rats with lesion of the medial amygdala do not have impaired salt appetite caused by salt depletion. Rats with lesion of the central nucleus of the amygdala or the bed nucleus of the stria terminalis (an area of brain intimately related to the central nucleus of the amygdala and often called the extended amygdala) show deficits in both deoxycorticosterone and salt-depletion-induced salt appetite. It seems clear at present that the several different brain areas mentioned earlier participate in the control of salt appetite. In rats, the subfornical organ and OVLT appear to be important to salt appetite caused by salt depletion, the medial amygdala seems to be important to salt appetite stimulated by the adrenocortical hormones, and the central nucleus of the amygdala and the bed nucleus of the stria terminalis seem to be involved in both. Further work is required to determine how these brain areas interact and what their roles are in modulating salt appetite during stress. Also, the brain mechanisms involved in salt appetite in other species still needs to be determined.

Conclusion

Overall, the data reviewed here show that stress can cause increased salt intake. Also, some, but not all, of the hormones associated with the stress response (e.g., ACTH and the adrenocortical hormones) can induce salt appetite. CRH and UCN inhibit salt appetite, even that caused by ACTH. The increased intake caused by ACTH appears to be mediated by the action of the adrenocortical hormones acting on receptors located in the amygdala and bed nucleus of the stria terminalis. At present, however, a role for adrenocortical hormones acting on receptors located in the circumventricular organs cannot be ruled out. An increase in angiotensin II or activity of the sympathetic nervous system caused by stress may also stimulate salt intake, possibly by the stimulation of the HPA

Table 1 Physiological effects of urocortin in sheep[a]

	CRH		*UCN*		*ACTH*		*ACTH + UCN*	
	ICV	*IV*	*ICV*	*IV*	*ICV*	*IV*	*ICV*	*IV*
Salt appetite	↓	↓	↓	↓	↓	↑↑↑	↓	↓
Corticosteroids[b]	↑	↑↑	↑	↑↑	↔	↑↑	↑↑	↑
MAP	↑	↑	↑	↑	↔	↑↑↑	↑↑↑	↓
Heart rate	↔	↔	↑↑↑	↑↑↑	↔	↑↑	↑↑	↑↑↑↑

[a]Arrows indicate direction and relative magnitude of change. ACTH, adrenocorticotropic hormone; CRH, corticotropin releasing hormone; ICV, intracerebroventricular; IV, intravenous; MAP, mean arterial pressure; UCN, urocortin.
[b]Plasma cortisol level.

axis. The mechanisms responsible for the inhibitory actions of CRH and UCN are as yet unknown.

Hormones associated with the stress response also increase blood pressure, and evidence suggests that the enhanced intake of salt may augment any increases in blood pressure caused by the stress. For example, in rats, increased blood pressure caused by stress is augmented by high salt intake. Similar observations have been made in dogs and baboons. In humans, salt sensitivity and family history may also contribute to the effect of high salt intake on stress-induced hypertension.

Figure 4 depicts some of the mechanisms by which stress may influence both dietary salt intake and blood pressure. Other factors will no doubt emerge with further research, including species differences.

Acknowledgments

This work was supported by the National Health and Medical Research Council of Australia (NHMRC grant number 217011) and the Deutsche Bank Postgraduate Scholarship.

Further Reading

Axelrod, J. and Reisine, T. D. (1984). Stress hormones: their interaction and regulation. *Science* **224**, 452–459.

Denton, D. A. (1982). *The hunger for salt: an anthropological, physiological and medical analysis*. Berlin: Springer-Verlag.

Denton, D. A., Coghlan, J. P., Fei, D. T., et al. (1984). Stress, ACTH, salt intake and high blood pressure. *Clinical and Experimental Hypertension: Theory and Practice* **A6**, 403–415.

Denton, D. A. and Nelson, J. F. (1980). Kare, M. R., Fregly, M. J. & Bernard, R. A. (eds.) *Biological and behavioral aspects of salt intake*, pp. 229–246. New York: Academic Press.

Gaillard, R. C. and Al-Damluji, S. (1987). Stress and the pituitary-adrenal axis. *Baillière's Clinical Endocrinology and Metabolism* **1**, 319–354.

Glavin, G. B. (1985). Stress and brain norepinephrine: a review. *Neuroscience and Biobehavioral Reviews* **9**, 233–243.

Makara, G. B. (1985). Mechanisms by which stressful stimuli activate the pituitary-adrenal system. *Federation Proceedings* **44**, 149–153.

Saavedra, J. M. (1992). Brain and pituitary angiotensin. *Endocrine Reviews* **13**, 329–380.

Schulkin, J. (1991). Sodium hunger: the search for a salty taste. Cambridge, UK: Cambridge University Press.

Weisinger, R. S., Blair-West, J. R., Burns, P., et al. (1996). The role of angiotensin II in ingestive behaviour: a brief review of angiotensin II, thirst and Na appetite. *Regulatory Peptides* **66**, 73–81.

Weisinger, R. S., Blair-West, J. R., Burns, P., et al. (2000). The inhibitory effect of hormones associated with stress on Na appetite of sheep. *Proceedings of the National Academy of Sciences USA* **97**(6), 2922–2927.

Weisinger, R. S., Blair-West, J. R., Burns, P., et al. (2004). Cardiovascular effects of long-term central and peripheral administration of urocortin, corticotropin-releasing factor, and adrenocorticotropin in sheep. *Endocrinology* **145**(12), 5598–5604.

Schizophrenia

M van den Buuse and D Copolov
The Mental Health Research Institute of Victoria, Parkville, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J M Crook and D L Copolov, volume 3, pp 393–397, © 2000, Elsevier Inc.

Introduction
Schizophrenia: Definition and Etiology
The Effects of External Stressors: Expressed Emotion and Life Events
The Hypothalamic-Pituitary-Adrenocortical (HPA) Axis and Schizophrenia
Stress, Early Development, and Schizophrenia: Neurobiological Findings
Treatment Implications of the Vulnerability-Stress Model of Schizophrenia: Psychosocial Treatment Strategies
Conclusion

Glossary

Adrenocorticotropin (ACTH)	The pituitary hormone that stimulates the synthesis and secretion of adrenal corticosteroids and especially the glucocorticoids.
Corticotropin-releasing hormone (CRH)	The 41-amino-acid residue peptide released from hypothalamic neurons that mediates neural control of ACTH secretion. CRH has also been implicated in the central control of emotion and behavior.
Dexamethasone suppression test (DST)	A test designed to determine the responsiveness of the HPA axis to the negative feedback effects of the potent synthetic glucocorticoid dexamethasone. Major mental disorder is often associated with a blunted response to the DST.
Expressed emotion (EE)	Emotion expressed by close relatives toward family members with schizophrenia.
Hypothalamic-pituitary-adrenocortical (HPA) axis	The system that involves CRH-induced ACTH release, which in turn leads to adrenal glucocorticoid secretion. The latter moderates ACTH release by way of negative feedback regulation.
Magnetic resonance imaging (MRI)	An imaging technique commonly used primarily to produce high-quality images of brain.
Ultra-high risk (UHR)	A cohort of individuals identified as being at ultra-high risk to develop psychosis.

Introduction

Stress may have an impact on the course and severity of schizophrenia during a number of stages along the neurodevelopmental trajectory, including presymptomatic and early symptomatic phases as well as the phases associated with chronicity and relapse. Extensive neurobiological research has now identified molecular pathways and brain regions that are likely to be involved in the interplay between stress and schizophrenia. The interaction between these pathways has not, however, been the subject of pharmacological intervention studies. Most of the therapy-oriented research in this field has involved psychosocial approaches to reducing stress and minimizing the impact of life events in order to reduce symptom severity and the risk of relapse.

Schizophrenia: Definition and Etiology

Schizophrenia is an enigmatic illness of unknown cause, heterogeneous symptomatology, and unpredictable course, with a variable treatment response. While the precise definition of schizophrenia poses numerous difficulties, there is general agreement that the disorder involves a breakdown of integration among emotions, thoughts, and actions due to underlying neurobiological abnormalities. Common symptoms include disorganized thought processes, hallucinations, delusions, inappropriate or blunted affect, social withdrawal, and poverty of speech. Cognitive impairments, including deficits in working memory, attention, and executive functioning are also often found in the disorder. Because many of these symptoms occur in other disorders, they cannot be regarded as diagnostic when taken individually. According to most diagnostic systems, the diagnosis of schizophrenia requires specific psychopathologies to be present for at least 6 months.

The onset of schizophrenia is usually during late adolescence or early adulthood. Gender differences exist, with women on average having a later onset and a more benign course of the illness. The lifetime prevalence of schizophrenia is around 1% of the general population across all ethnic backgrounds, nationalities, and socioeconomic classes. Most people with the disorder are unemployed. Many live in substandard

accommodation and experience drug or alcohol dependencies. Approximately 10% of people with schizophrenia commit suicide. Thus, schizophrenia represents a significant public health problem, from both an economic and a social point of view.

The Effects of External Stressors: Expressed Emotion and Life Events

Genetic factors play a major role in the development of schizophrenia; however, they alone do not explain the illness. First, multiple genes appear to be involved in the etiology of schizophrenia, collectively increasing susceptibility to the illness but each contributing only a small percentage to the risk. Furthermore, while family, twin, and adoption studies strongly support this genetic contribution, the finding of only 35–50% concordance for schizophrenia among monozygotic twins clearly supports the involvement of environmental factors. Epidemiological studies have identified several such adverse factors that may increase incidence of schizophrenia, such as pre- or perinatal obstetric complications, maternal influenza or rubella infection, or malnutrition and psychological trauma during pregnancy. For example, among both men and women conceived at the height of the Dutch famine in 1945, the risk of later developing schizophrenia was double that of individuals born before or after this period. Prenatal exposure to famine was also associated with decreased intracranial volume and white matter hyperintensities. These data have now been confirmed in a separate Chinese study, in which it was found that prenatal exposure to the 1959–61 famine more than doubled the risk of later developing schizophrenia. Both studies thus show that early stress caused by prenatal nutritional deficiency may play a role in the pathogenesis of schizophrenia. It should be noted that other nonstress factors could also play a role in the apparent link between famine and schizophrenia.

In addition to early environmental disturbances causing an increased risk, environmental factors later in life, such as drug abuse or psychosocial stress, may precipitate the disorder in predisposed individuals. For example, there is a significant relationship between the use of drugs of abuse, such as cannabis, and the exacerbation of psychosis and schizophrenia. The common denominator of all these early and late environmental factors is that they all induce stress, i.e., increased central and systemic release of stress hormones and disturbance of physiological and neurochemical homeostasis. The development of schizophrenia therefore may involve a disruption of early development, by either genetic or stress factors, which induces subtle changes along the neurodevelopmental trajectory. These changes could take place in synaptogenesis, synaptic pruning, axonal myelination, and neuronal proliferation in various brain regions, including cortex and hippocampus, but if they precede the onset of the disorder they either remain silent until the onset of the disorder or lead to very minor neuropsychiatric or behavioral changes in childhood and adolescence. Such changes have been reported and include a tendency for delayed neurological development, reduced social functioning, lower IQ, and neurological soft signs. However, while these associations are statistically significant, many individuals in the normal population also display such abnormalities. As a result, the presence of such mild abnormalities is not of clinical utility in the prediction of the development of schizophrenia. The emergence of psychotic symptoms may be triggered by additional stress, which may also be responsible for a later aggravation or reemergence of symptoms. One of the most influential versions of the neurodevelopmental hypothesis of schizophrenia thus comprises at least two principal components, highlighting postulated early changes in the brain, followed by later disturbances in psychological and physiological homeostasis – the so-called two-hit hypothesis of schizophrenia.

Early attempts to characterize the existence of stress-reactive schizophrenia were driven by the finding that attempts to overcome negative symptoms by social stimulation sometimes induced the reemergence of positive, psychotic symptoms of schizophrenia. This has since been tested through the study of expressed emotion (EE) – emotion expressed by close relatives toward family members with schizophrenia. Many, but not all, studies have shown that patients who experience a family environment with high EE, especially of a type that is chronic and critical in nature, have a greater propensity for psychotic relapse compared to patients from families with lower levels of EE.

Although EE studies have provided some of the strongest evidence for the stress reactivity of schizophrenia, the EE measure has been frequently misunderstood and misapplied. One of the most common mistakes is to conclude that critical and overinvolved caregivers are the major cause of poor clinical outcomes, with the result that family members are blamed for causing or exacerbating the illness. Such an approach fails to take into account the iterative relationship between high EE and patient relapse, which includes the impact of psychiatric illness on

the patient's carers. In support of the iterative nature of this relationship is the established correlation between EE and family burden.

In comparison to EE, which is relatively ambient and consistent, thereby providing a probe for the impact of chronic stress on schizophrenia, life events are more discrete and less predictable, thereby providing a greater insight into the effect of acute stress on the condition. Although the majority of studies support an association between the propensity to experience episodes of schizophrenia and either the frequency of both major life events and daily hassles or their negative subjective appraisal, most of these studies have been retrospective and could therefore be affected by a search after meaning, which leads individuals to bias their recall in favor of giving undue importance to life events in order to explain their psychotic symptoms.

The anxiety associated with the psychotic experience itself may also contribute to relapse and symptom severity. Clearly, responses to delusions of persecution and danger or to hallucinations of a derogatory and intrusive nature may involve severe distress, which in turn may further exacerbate psychotic symptoms.

The Hypothalamic-Pituitary-Adrenocortical (HPA) Axis and Schizophrenia

There is good evidence that HPA axis dysfunction, as evidenced by nonsuppression in the dexamethasone suppression test (DST), occurs in approximately one in five subjects with schizophrenia, which is more than in normal subjects but less than in patients with depression. Post mortem studies have shown reduced glucocorticoid receptor expression in the hippocampus of subjects with schizophrenia, which might contribute to reduced negative feedback to the HPA axis and thus play a role in DST nonsuppression.

DST nonsuppression appears to be more common in patients who are in the acute or initial phases of their illness than in stable or chronic phases. This may explain some inconsistency in the literature, which could be attributable to differences between prodromal, first episode, and chronic patients. In addition, antipsychotic treatment has been shown to normalize at least some of the HPA axis disturbances, such as nonsuppression in the DST. The initial phase of psychosis has been shown to be associated with HPA axis hyperactivity, shown by elevated circulating cortisol cerebrospinal fluid CRH levels and impaired DST. First episode patients show enlarged pituitary volumes as measured by magnetic resonance imaging (MRI). This might reflect HPA axis hyperactivity, since enhanced CRH and adrenocorticotropin (ACTH) secretion leads to increased size and number of pituitary ACTH-producing cells. These effects could be the consequence of the previously mentioned distressing psychotic experience or, alternatively, could be involved in the induction of psychosis. To distinguish between these possibilities, pituitary volumes have been measured in a cohort of individuals identified as being at ultra-high risk (UHR) to develop psychosis. Importantly, these subjects were symptom-free and did not receive antipsychotic medication at the time of pituitary volume measurement. Within this UHR cohort, pituitary volume positively predicted future transition to psychosis, and subjects with the greatest increase in pituitary volume also showed the shortest time between being scanned and psychosis onset. Theoretically, UHR individuals with enlarged pituitary volumes might benefit from preventative treatments that target HPA axis overactivity, such as glucocorticoid receptor antagonists or CRH receptor antagonists.

While there is clear evidence for disturbances in HPA axis function in a subset of individuals with schizophrenia, the mechanism by which these disturbances mediate illness development and symptomatology is less clear. Chronically enhanced levels of glucocorticoid hormones are neurotoxic, for example, leading to macroscopic volume reduction and cell loss in the hippocampus as well as reduced neuronal volume and dendritic arborization. These effects could be associated with some of the cognitive impairments seen in patients with schizophrenia and could also explain some of the structural changes seen in the brains of patients with schizophrenia, such as reduced cortical and hippocampal volumes.

There appears to be a synergistic relationship between HPA axis activity and dopaminergic activity in the brain. Glucocorticoids increase dopamine metabolism in the striatum and dopamine release in the nucleus accumbens. If these effects apply to all areas of the brain, HPA axis disinhibition is likely to contribute to dopaminergic hyperactivity, which appears to be central to psychosis symptomatology. Positron emission tomography studies have shown that patients with schizophrenia display greater increases in central dopamine release in response to certain challenges, including low-dose amphetamine. All antipsychotic drugs share dopamine receptor antagonism as part of their pharmacological mode of action. Conversely, abnormalities in dopaminergic activity influence HPA axis activity and render an individual hyperresponsive to stress. The vulnerability-stress model of schizophrenia suggests that the illness is potentiated by stress due to this interplay with brain dopaminergic activity.

Stress, Early Development, and Schizophrenia: Neurobiological Findings

In an attempt to elucidate neurobiological mechanisms involved in the interaction of environmental factors and schizophrenia, animal model studies have simulated major life events during early stages of development. These animals, mostly rats, are then tested as adults in behavioral paradigms with relevance to schizophrenia. Developmental disruptions include neonatal stress induced by maternal deprivation, prenatal maternal stress, developmental stress induced by placental insufficiency, antimitotic treatment or vitamin D depletion, and prenatal or neonatal infection. Several studies have also used neonatal brain lesions, such as in the ventral hippocampus, frontal cortex, or amygdala. Such early interventions induce behavioral effects in the adult such as hyperlocomotion, enhanced sensitivity to psychotomimetic drugs (amphetamine, phencyclidine), disruptions of prepulse inhibition, and learning and memory deficits.

Another approach has been to study the influence of variations in maternal care. In these rat studies, dams were classified according to the level of maternal care, and behavioral and molecular changes in the offspring were examined. Offspring from mothers that displayed high levels of pup licking, grooming, and nursing differed in the level of DNA methylation from offspring from mothers that displayed low levels of these behaviors. These effects persisted into adulthood and resulted in specific changes in the regulation of glucocorticoid receptor expression. Offspring from mothers that displayed lower levels of maternal care showed lower levels of glucocorticoid receptor expression and greater plasma corticosterone increases to stress than offspring from mothers that displayed higher levels of maternal care. These studies show that early environmental factors, in this case the level of maternal care, induce long-term molecular and biochemical alterations leading to differences in glucocorticoid receptor expression and alterations in HPA axis activity. Such changes in stress responsiveness could be involved in the role of stress in schizophrenia.

Treatment Implications of the Vulnerability-Stress Model of Schizophrenia: Psychosocial Treatment Strategies

Psychosocial treatments are critical elements in the management of schizophrenia, provided that they are – as they need to be – in tandem with the keystone of treatment, antipsychotic medication. Stress management has received relatively little research attention in the treatment of the disorder, but there are suggestions that this form of treatment reduces the prevalence of hospitalization. Psychosocial treatment modalities that have received greater validation include family therapy, assertive community treatment, psychoeducation, social skills training, and cognitive-behavioral therapy. To the extent that these modalities improve coping skills, the positive outcomes they have been associated with in terms of relapse prevention and symptom amelioration might, in part, be a consequence of reduction in stress levels.

Conclusion

A variety of epidemiological, clinical, and neurobiological studies have suggested a nontrivial link between stress and schizophrenia, both during development and during the initial and established phases of the disease. These studies provide clues about brain mechanisms that could be the target of future new pharmacological treatments and about psychosocial strategies aimed at enhancing coping skills and reducing the impact of stress on the individual. However, as the available evidence is still controversial, further studies are clearly needed to better our understanding of the role of stress in schizophrenia.

See Also the Following Articles

Dexamethasone Suppression Test (DST); Hypothalamic-Pituitary-Adrenal; Pre-pulse Inhibition; Psychotic Disorders.

Further Reading

Bustillo, J., Lauriello, J., Horan, W. and Keith, S. (2001). The psychosocial treatment of schizophrenia: an update. *American Journal of Psychiatry* **158**, 163–175.

Corcoran, C., Walker, E., Huot, R., et al. (2003). The stress cascade and schizophrenia: etiology and onset. *Schizophrenia Bulletin* **29**, 671–692.

Cotter, D. and Pariante, C. M. (2002). Stress and the progression of the developmental hypothesis of schizophrenia. *British Journal of Psychiatry* **181**, 363–365.

Garner, B., Pariante, C. M., Wood, S. J., et al. (2005). Pituitary volume predicts future transition to psychosis in individuals at ultra-high risk of developing psychosis. *Biological Psychiatry* **58**, 417–423.

Gispen-de Wied, C. C. (2000). Stress in schizophrenia: an integrative view. *European Journal of Pharmacology* **405**, 375–384.

Hulshoff Pol, H. E., Hoek, H. W., Susser, E., et al. (2000). Prenatal exposure to famine and brain morphology in schizophrenia. *American Journal of Psychiatry* **157**, 1170–1172.

McGrath, J. J., Feron, F. P., Burne, T. H., Mackay-Sim, A. and Eyles, D. W. (2003). The neurodevelopmental hypothesis of schizophrenia: a review of recent developments. *Annals of Medicine* **35**, 86–93.

Miller, P., Lawrie, S. M., Hodges, A., et al. (2001). Genetic liability, illicit drug use, life stress and psychotic symptoms: preliminary findings from the Edinburgh study of people at high risk for schizophrenia. *Social Psychiatry and Psychiatric Epidemiology* **36**, 338–342.

Mueser, K. T. and McGurk, S. R. (2004). Schizophrenia. *The Lancet* **363**, 2063–2072.

Pantelis, C., Yucel, M., Wood, S. J., McGorry, P. D. and Velakoulis, D. (2003). Early and late neurodevelopmental disturbances in schizophrenia and their functional consequences. *Australia and New Zealand Journal of Psychiatry* **37**, 399–406.

Van den Buuse, M., Garner, B. and Koch, M. (2003). Neurodevelopmental models of schizophrenia: effects on prepulse inhibition. *Current Molecular Medicine* **3**, 459–471.

Van den Buuse, M., Garner, B., Gogos, A. and Kusljic, S. (2005). Importance of animal models in schizophrenia research. *Australia and New Zealand Journal of Psychiatry* **39**, 550–557.

Weaver, I. C., Cervoni, N., Champagne, F. A., et al. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience* Aug **7**(8), 847–854.

School Stress and School Refusal Behavior

C A Kearney, L C Cook and G Chapman
University of Nevada, Las Vegas, NV, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C A Kearney and C Pursell, volume 3, pp 398–401, © 2000, Elsevier Inc.

School Stress and Linkage to School Refusal Behavior
Major Characteristics of School Refusal Behavior
Etiology and Major Functions of School Refusal Behavior
Assessment of School Refusal Behavior
Treatment of School Refusal Behavior

Glossary

School refusal behavior	The substantial child-motivated refusal to attend school and/or difficulties remaining in classes for an entire day.
School stress	Unpleasant physical and cognitive symptoms in response to global and specific school-related stressors.

School Stress and Linkage to School Refusal Behavior

Children and adolescents become anxious about many things, but youths become particularly stressed by difficult family and school situations. School-related stressors may involve global and specific issues. Global school-related stressors include large class size, threats of victimization and violence, changes in academic and social status, excessive homework, and tedious curricula. Specific school-related stressors include teachers and other school officials, peer interactions, examinations, performances before others, social gatherings, and use of cafeterias and public restrooms.

School-related stressors are quite prevalent among youths, and many youths experience concomitant physiological and cognitive symptoms as a result. Common physiological symptoms include headaches, stomachaches, shortness of breath, trembling, dizziness, and heart palpitations. Common cognitive symptoms include catastrophic thoughts about losing control, going insane, suffering humiliation or harm, experiencing negative evaluation from others, appearing foolish, having few friends, having trouble concentrating, and lacking competence in various areas. If school-related stress becomes severe, then some youths may refuse to attend school to avoid these unpleasant symptoms.

Major Characteristics of School Refusal Behavior

School refusal behavior is the substantial child-motivated refusal to attend school and/or difficulties remaining in classes for an entire day. Specifically, the behavior involves youths ages 5–17 years who (1) are completely absent from school, (2) attend school but then leave school during the school day, (3) go to school following intense behavior problems such as tantrums in the morning, and/or (4) display unusual distress during the school day that leads to pleas for future nonattendance. School refusal behavior may also be defined by its timeline. For example,

self-corrective school refusal behavior refers to absenteeism that remits spontaneously within 2 weeks. Acute school refusal behavior lasts from 2 weeks to 1 year, and chronic school refusal behavior lasts longer than 1 year. Greater chronicity is related to greater difficulty reintroducing a child to school.

Approximately 5–28% of school-age youths refuse school at some time in their lives. School refusal behavior is seen equally in boys and girls and among families of various income levels. The short-term consequences of school refusal behavior include declining academic status, social alienation, increased risk of legal trouble, family conflict, and severe disruption in a family's daily routine. The long-term consequences of school refusal behavior include school dropout and subsequent economic deprivation, occupational and marital problems, alcohol abuse and criminal behavior, and mental disorders.

School refusal behavior is marked by considerable symptom heterogeneity. Common internalizing problems include general and social anxiety/shyness, depression and social withdrawal, fear, fatigue, and somatic complaints such as stomachaches, headaches, nausea, and tremors. Common externalizing problems include tantrums, verbal and physical aggression, clinging, reassurance seeking, refusal to move, noncompliance, and running away from school or home.

School refusal behavior is also marked by many different diagnoses that could cause or be the result of problematic absenteeism. In this population, common comorbid diagnoses include separation anxiety disorder, generalized anxiety disorder, depression, and oppositional defiant disorder, among others. In addition, school refusal behavior has been linked to limited developmental disabilities such as learning disorders and pervasive developmental disabilities such as Asperger's disorder, autism, and mental retardation.

Etiology and Major Functions of School Refusal Behavior

The specific causes of school refusal behavior are often unclear for a particular case. However, common triggers include the onset of a new school year, entry into a new school building, school-related trauma, teacher or class changes, family illness or transition, parental psychopathology, and marital conflict or divorce. Child factors such as behavioral inhibition, excessive shyness, and association with deviant peers can also be influential. Unfortunately, identifying the etiology in this population is often a fruitless task. As such, clinicians may focus instead on the primary functions or reasons why youths refuse school:

1. To avoid school-related objects and situations that provoke a general sense of negative affectivity (dread, anxiety, depression, and physiological symptoms).
2. To escape aversive social and/or evaluative situations at school.
3. To receive attention from significant others outside of school.
4. To obtain tangible reinforcement outside of school.

The initial two functions refer to youths who refuse school for negative reinforcement, or to avoid something unpleasant at school. These functions are closely related to the concept of school stress. The latter two functions refer to youths who refuse school for positive reinforcement, or to pursue someone or something more alluring outside school. Some youths also refuse school for multiple reasons, and those who do often require complex and extended treatment.

Assessment of School Refusal Behavior

Assessing youths with school refusal behavior often involves defining a child's exact behavior problems, discovering primary and secondary functions for absenteeism, and identifying effective treatment. To do so, assessment methods may include interviews, child self-report and parent/teacher measures, direct observation, daily monitoring, and discussions with school officials. Interviews may be structured or unstructured in nature. A common structured interview for this population is the Anxiety Disorders Interview Schedule for Children, which has child and parent versions and helps clinicians identify various diagnoses. In addition, the interview contains a special section on school refusal behavior to help clinicians pinpoint the frequency and intensity of the behavior and related stress-based symptoms. Unstructured interviews may also be used to assess youths with school refusal behavior. Questions may relate to school-related stressors that affect a particular child, degree of avoidance, physiological and cognitive symptoms, attention-seeking, and pursuit of tangible reinforcers outside of school.

Several child self-report and parent and teacher measures also pertain to youths with school refusal behavior. The most common address negative affectivity, depression, general and social anxiety, somatic complaints, overactivity, aggression, noncompliance, demands for attention, social problems, fear, self-concept, family dynamics, and the function of school refusal behavior. Gathering information from multiple sources is recommended because school refusal behavior can include many subtle symptoms.

Direct observation of a child's and family's morning activities at home may also help define the primary symptoms and maintaining variables of school refusal behavior. Pertinent behaviors include:

1. Avoidance in the form of clinging, refusal to move, running away, and/or noncompliance to requests.
2. Physiological reactivity such as stomachaches, headaches, abdominal pain, tremors, and nausea/vomiting.
3. Cognitive distortions or verbalizations about discomfort related to school.
4. Sudden changes in behavior.
5. Pleas to stay home from school.
6. Increased family conflict, especially following a curtailment of the child's social activities.
7. Significant changes in parental behavior.
8. Teacher reports of differences in a child's behavior at school.

The daily assessment of a child's school refusal behaviors and attendance is also important. This helps gauge familial compliance to homework assignments, increase insight into a child's problem, and identify positive or negative changes in behavior across time. For example, family members may provide daily ratings with respect to a child's anxiety, depression, and overall distress (general feelings of dread or being upset). In addition, parents may rate a child's noncompliance (not listening to parent commands) and disruption to their daily routine. Parents can also list other child behavior problems and time missed from school.

Finally, school officials may be contacted for additional information, including course schedules, grades, written work, and required make-up work; goals and attitudes of school officials and peers regarding a child; procedures and timelines for reintegrating a child into school; potential obstacles to reintegrating a child into school; confirmation of past school refusal behavior; general social or other behaviors of a child in school; a map or outline of the school (e.g., lockers, cafeteria, library); feedback about treatment effectiveness; disciplinary procedures and procedures for contacting parents; and rules about absenteeism, conduct, or leaving school areas.

Treatment of School Refusal Behavior

Treatment for youths with school refusal behavior may be linked to the function of the behavior. For example, treatment for youths who refuse school to avoid stimuli provoking negative affectivity (i.e., school stress) often involves reducing unpleasant physiological symptoms and exposing a child to various school-related items and situations. Specifically, this involves psychoeducation about the nature and process of anxiety (i.e., linking feelings, thoughts and behaviors), developing an anxiety and avoidance hierarchy of situations and stimuli, teaching somatic management skills to decrease negative emotional arousal (i.e., relaxation training and breathing retraining), systematic imaginal and *in vivo* exposure to anxiety cues identified on a graded hierarchy (progressive reintroduction to the school setting), and having a child praise him- or herself for coping with transient negative emotions.

For youths who refuse school to escape aversive social and/or evaluative situations, treatment strategies involve building or refining social skills and reducing social anxiety in key interactive situations at school. Specifically, this involves psychoeducation about the nature and process of anxiety (i.e., linking feelings, thoughts, and behaviors), teaching a child to identify cognitive distortions (e.g., all-or-none thinking, catastrophization, and overgeneralization) and negative self-statements in anxiety-provoking situations, learning methods to change distorted thoughts to helpful coping thoughts, graduated exposure to anxiety-provoking social situations through in-session practice with a therapist (including modeling and role-play of different situations), and between-session practice of skills in real-life social situations.

For youths who refuse school for attention, treatment strategies involve parent-based techniques to shift parental attention away from school refusal behaviors and toward appropriate school attendance behaviors. Specifically, this involves restructuring parent commands, establishing fixed routines regarding a child throughout the day, implementing negative consequences for school refusal behaviors, implementing positive consequences for school attendance behaviors, and, in some cases, forced school attendance.

For youths who refuse school for tangible reinforcement outside school, treatment strategies involve improving a family's problem-solving abilities, reducing conflict, increasing rewards for school attendance, and decreasing rewards for school absenteeism. Specifically, this involves establishing times and places when family members can negotiate problem solutions, defining behavior problems specifically, designing written parent–child contracts to address problems, implementing contracts, training in communication skills, training in peer-refusal skills (i.e., helping a child refuse offers to miss school), and escorting a child to school or to each class. For all treatments, relapse-prevention training is also necessary.

See Also the Following Articles

Attention-Deficit/Hyperactivity Disorder, Stress and; Neurodevelopmental Disorders in Children.

Further Reading

Berg, I. and Nursten, J. (1996). *Unwillingly to school* (4th edn.). Washington, DC: American Psychiatric Press.

Kearney, C. A. (2001). *School refusal behavior in youth: a functional approach to assessment and treatment*. Washington, DC: American Psychological Association.

Kearney, C. A. and Albano, A. M. (2000). *When children refuse school: a cognitive-behavioral therapy approach/ therapist's guide*. New York: Oxford.

Kearney, C. A. and Silverman, W. K. (1996). The evolution and reconciliation of taxonomic strategies for school refusal behavior. *Clinical Psychology: Science and Practice* **3**, 339–354.

Kearney, C. A. and Silverman, W. K. (1999). Functionally-based prescriptive and nonprescriptive treatment for children and adolescents with school refusal behavior. *Behavior Therapy* **30**, 673–695.

King, N. J., Ollendick, T. H. and Tonge, B. J. (1995). *School refusal: assessment and treatment*. Boston, MA: Allyn and Bacon.

Silverman, W. K. and Albano, A. M. (1996). *Anxiety disorders interview schedule for DSM-IV: child version*. New York: Oxford.

School Violence and Bullying

J Juvonen
University of California, Los Angeles, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Hostile School Behavior: Prevalence and Concerns
Connection between Bullying and Violent Offending
Antecedents and Consequences of Hostile Experiences
The Effects of Witnessing Hostile Behaviors

Bullying and school violence are significant sources of stress in American schools. In a recent national survey, children ages 8–15 rated bullying as one of the major concerns affecting their lives. Another survey indicated that almost two-thirds of public secondary school students thought that a shooting could take place in their school.

In contrast to violent behaviors that involve use of physical force, bullying entails the abuse of psychological strength, which can be achieved by various means. Bullies intimidate their victims by relying on name-calling, exclusion, threats, and/or spreading of rumors. Because the perpetrators of violent acts as well as bullying abuse their power, both forms of hostile behavior cause psychological distress in that the actual and potential victims of peer maltreatment or abuse experience anxiety. Repeated experiences of hostile behavior by schoolmates are, in turn, associated with increased feelings of depression and either withdrawn or aggressive behavior.

Hostile School Behavior: Prevalence and Concerns

According to the most recent national data, 14% of students ages 12–18 reported having been bullied in school during the previous 6 months. The estimates for bullying vary depending on how broadly the term is defined and the time frame provided. Using daily (as opposed to monthly or yearly) reports of bullying experiences, almost 50% of young teens reported at least one incident of bullying during a 2-week period in urban middle schools. Up to 70% of secondary students reported having experienced bullying at some point during their school careers. Although the experiences of bullying are fleeting for most students, an estimated 5–10% of students confront bullying on a weekly, if not daily, basis for 1 or more school years.

The prevalence estimates of bullying also vary by grade level and school type. Compared to other grade levels, middle school students reported the highest rates of bullying, and sixth-grade students were most fearful of their safety. Although girls reported being more concerned about their safety at school or on their way to and from school, boys reported more incidents of violence and bullying. Boys were also the perpetrators of the most violent incidents.

Approximately 70% of public schools reported one or more violent incidents during a school year. Although the most serious forms of violence are less common, one-fifth of public schools reported at least one incident involving aggravated assault, sexual assault, rape, or robbery per year. Violent incidents

are most likely to take place during the high school years and in large urban schools.

Students in urban public schools are most likely to fear for their safety. Black and Latino students reported feeling more fearful of being attacked than did White students. Although some hostile incidents appeared to be racially motivated, this does not necessarily mean that students were worst off in ethnically mixed schools. Recent analyses of middle schools suggest that greater ethnic diversity (i.e., a larger number of ethnic groups and their relative balance) is associated with feelings of safety.

Connection between Bullying and Violent Offending

Children who bully in childhood are at risk of becoming violent offenders. Research conducted in Norway shows that boys identified as bullies in their teenage years were four times as likely to commit violent acts in their twenties than other boys. Long-term studies conducted in the United States reveal that when childhood bullying is associated with more serious subsequent fighting, the risk of committing assaults and rape significantly increases. Hence, at least the physical forms of bullying and violent offending among males reflect a spiral of increasingly serious hostile behaviors.

Both childhood bullying and school violence are predicted by personality dispositions (e.g., impulsivity), environmental factors beyond school (e.g., problematic relationships with parents), and school-related situations (e.g., perceptions of unfair treatment by teachers or peers). These risk factors often interact in that students with certain dispositions are more likely to encounter situations that prompt them to act aggressively. For example, impulsive students are likely to act in a way that increases their chances of someone's bumping into them in a hallway, and, after experiencing an unintended push, they are more likely to react with hostility. The hostile response is often promoted by a perception that the other person intended to cause harm.

Many aggressive students develop a generalized sense of mistrust. Hence, students commit hostile acts to protect themselves. For example, compared to nonbullied students, bullied students are more likely to carry a weapon to school. Fear for their safety is the most common reason provided by students who carry weapons to school.

Other students' hostile acts are motivated by a desire to maintain a sense of control and to show their power. These bullies or violent offenders enjoy at least temporary admiration among their peers because of their dominant status. Fellow students rarely challenge their mean and hurtful behaviors because they are fearful of the reactions of these dominant ringleaders. Passivity and acceptance of toughness, in turn, promote social norms that encourage and maintain bullying or even violent behavior. Thus, the school environment, including the reactions of fellow peers (victims or bystanders), plays a critical role in motivating hostile behavior in school.

Antecedents and Consequences of Hostile Experiences

Because of the low prevalence rates of serious violent incidents in school, little is known about the factors that increase the risk of being targeted. However, research on bullying shows that socially withdrawn and passive children are at risk of getting bullied. After repeated experiences of bullying, these submissive victims become even more withdrawn. These students most often suffer in silence, and teachers and parents are unaware of their plight or distress. Instructors may interpret the quiet signs of distress as a lack of interest in schoolwork. Indeed, teachers rate victims as being just as disengaged from schoolwork as bullies.

Students who respond to bullying with unregulated aggression (i.e., aggressive victims) are also at a high risk for continued peer maltreatment. In these cases, bullying experiences are related to extreme peer rejection and severe school difficulties. Aggressive victims are presumed to have underlying emotional difficulties, which are also used to help explain why their difficulties continue. Thus, in the cases of both submissive and aggressive victims, the causes and consequences of bullying form a vicious circle.

Although aggressive victims resemble violent offenders most closely, the victims of bullying are unlikely to resort to violence to get back at their tormentors. Nevertheless, the tragic school shooting incidents since the late 1990s highlight that chronic victims of bullying, like violent offenders, have trouble effectively dealing with social failures (e.g., public ridicule or rejection by a romantic partner). Some violence prevention and antibullying interventions are therefore specifically designed to help students better cope with bullying incidents.

The Effects of Witnessing Hostile Behaviors

The hostile behaviors of schoolmates affect not only those who are victimized but also those who witness peer abuse or maltreatment in school. For example, analyses of daily incidents of bullying in middle school show that social anxiety is increased when students see someone else being bullied. Students

who witness violent incidents are likely to feel especially vulnerable and unsafe. Thus, although there may be relatively few hostile students in any one school, one or two can terrorize entire grade levels and a couple of violent events may cause widespread fear. In view of such exposure effects, the level of concern about bullying and violence in American schools is not unrealistically high.

Part of the distress caused by observed events is likely to be related to the school's response system (or lack thereof). Although the parents' concerns about the safety of their children at school may be alleviated by the presence of metal detectors and campus officers, students continue to worry about their safety because they know that someone can smuggle a weapon to campus or they are fearful of someone bullying them in sites lacking monitoring (e.g., restrooms and locker rooms). In addition to feeling vulnerable, an unsafe or noncaring school climate is likely to breed mistrust and alienation that, in turn, increase the risk of students carrying weapons and setting off a hostile response.

In summary, school violence and bullying cause, and are also partly caused by, feelings of vulnerability that reach beyond personal experiences of hostile acts. This can be depicted as a following process:

Perceived vulnerability of self → Hostile response to perceived threat → Unsafe school climate

To promote feelings of safety and trust, schools are relying more and more on prevention models that focus not only on at-risk youth but also on helping improve the school climate. A caring climate in schools, in which students feel they belong, are respected, and are listened to and in which teachers help mediate hostile incidents, can buffer the safety concerns and distress of the entire student body.

See Also the Following Articles

Adjustment Disorders; Depression Models; School Stress and School Refusal Behavior.

Further Reading

Jimerson, S. R. and Furlong, M. (eds.) (2006). *Handbook of school violence and school safety: from research to practice*. Mahwah, NJ: Lawrence Erlbaum.

Juvonen, J. and Graham, S. (eds.) (2001). *Peer harassment in school: the plight of the vulnerable and victimized*. New York: Guilford Press.

Juvonen, J., Nishina, A. and Graham, S. (2006). Ethnic diversity and perceptions of safety in urban middle schools. *Psychological Science* **17**, 393–400.

Leary, M. R., Kowalski, R. M., Smith, L., et al. (2003). Teasing, rejection, and violence: case studies of the school shootings. *Aggressive Behavior* **29**, 202–214.

Smith, P. K., Pepler, D. and Rigby, K. (eds.) (2004). *Bullying in schools: how successful can interventions be?* New York: Cambridge University Press.

Relevant Websites

U.S. Department of Education and U.S. Department of Justice. (2005). Indicators of school crime and safety: 2005. National Center of Education Statistics. http://nces.ed.gov.

Seasonal Changes in Stress Responses

R J Nelson and L B Martin II
Ohio State University, Columbus, OH, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R J Nelson and D L Drazen, volume 3, pp 402–408, © 2000, Elsevier Inc.

Introduction
Seasonal Energetic Adaptations
Seasonal Adjustments in Glucocorticoid Secretion
Why Are Stress Responses Seasonally Variable?

Glossary

Glucocorticoids	A class of steroid hormones secreted from the adrenal cortices that serves to promote energy use. Corticosterone is the glucocorticoid usually secreted by rodents and birds, whereas cortisol is typically secreted by primates.
Melatonin	An indole-amine hormone secreted from the pineal gland at night that is important in the regulation of daily and seasonal biological rhythms.
Photoperiod	Day length or the amount of light per day.

Photoperiodism	The response of plants to the length of day and night, including the ability to determine day length (usually by measuring night length) in both plants and animals; first proposed by botanists Garner and Allard in 1922.
Pineal gland	An endocrine gland that secretes melatonin, usually located between the telencephalon and diencephalon in vertebrates; also called the epiphysis.

Introduction

Stressors are not constant but, rather, vary over time. In addition to daily changes in stress, seasonal changes in the quality and quantity of stress are common. Many energetic adaptations have evolved that attenuate the stress response during the winter when energy shortages may limit the ability of organisms to cope successfully with specific types of stressors. The temporal organization of such stress responses therefore must be considered to obtain a holistic understanding of the mechanisms underlying the stress response.

The main endocrine features of the stress response involve hormones that act to suppress energy storage and promote energy utilization from adipose and liver stores. Epinephrine and the glucocorticoids are the major hormonal components of the stress response. Animals require a relatively steady supply of energy to sustain biological functions. Although these energy demands are somewhat continual, both daily and seasonal fluctuations in energy requirements occur as physiological and behavioral activities wax and wane. Animals typically do not feed continuously to supply energy to cells; rather, adaptations have evolved to allow animals to store and conserve energy when they are engaged in nonfeeding activities. Eventually, however, the chronic need for energy depletes these stores and requires animals to obtain food. In most habitats, significant daily and seasonal fluctuations in food availability occur. For example, outside of the tropics, food availability is generally low during the winter when thermogenic energy demands are typically high. Consequently, energy intake and energy expenditures are never perfectly balanced because of fluctuations in both energy requirements and availability. Whenever a significant imbalance of energy is perceived, a stress response in the form of secreted glucocorticoids ensues.

Reproduction is sufficiently demanding that a stress response is often associated with breeding in terms of elevated glucocorticoid concentrations. That is, glucocorticoids are released to support the energetic needs of reproduction including territorial defense or courtship behaviors. However, when energy is insufficient to support reproduction and thermogenesis or other energetic demands, the stress response becomes prolonged. A chronic stress response results in pathology. Most animals restrict breeding to specific times of the year when food is abundant and survival and reproductive success are most likely. The inhibition of winter breeding is one of the central components of a suite of energy-saving seasonal adaptations evolved among nontropical animals. In addition to the seasonal onset and cessation of reproductive function among nontropical vertebrates, it is also well-established that endocrine, metabolic, growth, neural, immune, and thermogenic processes undergo seasonal changes to promote winter energy conservation. Mechanisms in virtually every physiological and behavioral system have evolved to cope with the winter energetic bottleneck resulting from increased thermogenic energy requirements during periods of low energy availability. Presumably, animals possessing these energy-conserving adaptations enhance their survival and ultimately increase their reproductive success. Generally, glucocorticoid concentrations are higher during the winter than the summer and individuals display more elevated stress responses during the winter than the summer. However, because coping strategies are engaged in anticipation of the winter stressors, many individuals do not show exaggerated stress responses during the winter.

How do animals determine time of year? In most cases, the photoperiod (day length) acts as the principal external cue to determine time of year. Photoperiod measurement is important for organizing the seasonal collapse and regrowth of the reproductive system in mammals and birds from temperate and boreal latitudes. Indeed, changes in day length provide the most error-free signal for predicting favorable conditions. Day length is encoded by the hormone melatonin, which is secreted primarily at night from the pineal gland.

Seasonal Energetic Adaptations

Many aspects of life are energetically challenging, and individuals have evolved timing mechanisms to parse out energetically demanding activities to different times of year. For example, birds engage in predictable cycles of migration, breeding, parenting, molting, and overwintering. These energetically expensive activities generally do not overlap, and, again, all these activities elicit classical stress responses (i.e., elevated activation of the hypothalamic-pituitary-adrenal (HPA) axis).

One 1936 study is particularly informative in understanding the basis of Selye's thinking about stress

responses. Rats were housed in low-temperature conditions for 2 days, 2 weeks, or 2 months. After 2 days in these conditions, Selye noted the hypertrophy of the adrenal cortex; reduced eosinophil (a type of immune cell circulating in the blood) numbers; and atrophy of the thymus, spleen, and lymph nodes. In other words, the rats exhibited a classic stress response. After 2 weeks in the low temperatures, it was discovered that rats no longer displayed these symptoms; essentially, they had become cold-adapted and could tolerate low temperatures without exhibiting any further stress responses. After approximately 2 months of exposure to low temperatures, however, the rats lost this acquired resistance to low temperatures and died.

This process of coping with stressors was termed the general adaptation syndrome (GAS), and consists of three stages: (1) the alarm reaction, (2) resistance, and finally, (3) exhaustion. During the alarm reaction, stressors are detected. Coping with the stressor occurs during the resistance stage. The exhaustion stage is characterized by the termination of the stress response and the onset of stress pathology, which in extreme cases can lead to death. Although Selye believed that the exhaustion phase was due to the termination of the stress response, current ideas about stress indicate that the pathological effects of stress are due to the prolonged activation of the stress response. These three stages of the GAS correspond to the response of the rats to low temperatures after 2 days (alarm reaction), 2 weeks (resistance), and 2 months (exhaustion), respectively. If individuals have the opportunity to gradually move to low temperatures or are allowed to prepare for low-temperature conditions by first being housed for 8–10 weeks in short day lengths, then no stress responses occur.

Seasonal Adjustments in Glucocorticoid Secretion

Glucocorticoids are released in response to stressful stimuli and can affect reproduction, digestion, growth, multiple types of immune function, and virtually any other physiological process that requires significant energy. Predictably, seasonal variation in glucocorticosteroid secretion is more often the rule than the exception, probably because of the pivotal role these hormones play in orchestrating the distribution of resources among competing physiological activities.

Most animals studied to date display some seasonal variation in stress responses. For instance, Indian barnacles (*Balanus balanoides*) exhibit elevated antioxidant enzyme activity (a component of this species' stress response) prior to the monsoon season compared with during it. Further, almost all wild amphibians, reptiles, and birds studied to date show marked changes in glucocorticoid concentrations during the year, with maximal values observed during the breeding season. Individuals of most mammalian species exhibit similar basal glucocorticosteroid concentrations year-round in the wild. However, this apparent pattern may reflect the paucity of studies on wild rodents rather than a lack of seasonality of this parameter. Indeed, rodents, such as deer mice (*Peromyscus* sp.), bred and maintained in the lab exhibit high corticosteroid concentrations when kept in long-day (breeding) conditions. Other species, however, such as Siberian hamsters (*Phodopus sungorus*), display elevated glucocorticosteroid concentrations when housed in short photoperiods.

Although it is clear from these studies that stress responses are temporally labile, two questions remain: (1) Is seasonal variation omnipresent within and among species (e.g., over a species' range) and (2) are other aspects of the stress response as seasonally labile as corticosteroid hormone concentrations? In terms of spatial variability, the results are mixed. Presumably, if photoperiod has significant effects on stress responses, then animals living at higher latitudes would exhibit higher circulating corticosteroid concentrations during breeding than animals at low latitudes simply because of the large variation in annual photoperiod. Redpolls (*Carduelis flammea*) are birds that live in the challenging arctic conditions of Alaska. Both males and females showed the typical marked elevation in blood corticosterone concentrations when captured for approximately 1 h, indicating that they responded similarly to many other vertebrates. However, the amount of the elevation in postcapture corticosterone varied across the year; corticosterone concentrations were generally blunted during January and maximal during June when birds were breeding. Generally, birds that had the highest fat stores secreted the lowest amounts of glucocorticoids postcapture. Similar patterns have been reported for other birds living at high latitudes, including the white-crowned sparrows (*Zonotrichia leucophrys gambelii*), snow buntings (*Plectrophenax nivalis*), and Lapland longspurs (*Calcarius lapponicus*). Across latitudes, however, results are mixed. Some species show little variability in circulating glucocorticoid concentrations (e.g., pied flycatchers, *Ficedula hypoleuca,* and mountain chickadees, *Poecile atricapilla*), whereas others (e.g., house sparrows, *Passer domesticus,* and white-crowned sparrows, *Zonotrichia leucophrys*) exhibit spatial variation. In common with spatial variability, the social system of animals can also affect seasonal variation in glucocorticoid responses. A study of greylag geese (*Anser anser*)

revealed that social status and physiological stress are not related in the same way in all seasons of the year.

In addition to hormones, other aspects of the glucocorticoid stress response show significant spatial and temporal flexibility. In particular, the HPA axis appears to be substantially malleable among wild songbirds. Individuals of different species can apparently modulate the levels of glucocorticoids they produce at different times of the year by changing the sensitivity of adrenergic tissues to upstream regulatory hormones such as corticotropin releasing hormone (CRH) or adrenocorticotropic hormone (ACTH). Alternatively, some species are capable of modulating steroid-binding globulin levels in their blood across the year instead, although it remains unclear whether these globulins promote or inhibit the actions of the glucocorticosteroids to which they bind. Finally, corticosteroid receptors also appear to be seasonally variable. In house sparrows, membrane corticosteroid receptors are lowest during the nesting season, whereas cytosolic receptors are lowest during winter.

Why Are Stress Responses Seasonally Variable?

Presumably, stress responses help animals prepare for or recover from stressors. Why, then, are stress responses variable during the year when all individuals could benefit from optimal stress responses? Several hypotheses have been proposed. First, glucocorticoid stress responses are expensive and hence may not be affordable at all times of the year. Indeed, pyruvate, a readily usable energy source, blocked the stress-evoked suppression of cell-mediated and humoral immune activity in mice (*Mus musculus*). Second, high corticosterone concentrations, particularly when maintained for extended periods (i.e., chronic), can be immunosuppressive; over the short term, however, glucocorticosteroids can be immunoenhancing. In general, winter survival in small animals is expected to require a positive balance between short-day enhancement of immune defenses and glucocorticoid-induced immunosuppression. Immunosuppression caused by high glucocorticosteroids may be induced by many winter-related factors, including overcrowding, increased competition for scarce resources, low temperatures, reduced food availability, increased predator pressure, and lack of shelter. Moreover, winter breeding, with the potentially concomitant elevation in sex steroid hormones, may also compromise the immune system. Presumably, this explains why winter breeding predominantly occurs when other environmental stressors such as temperature and food availability are not severe. Overall, the balance between enhanced immune function (i.e., to the point where autoimmune disease becomes a danger) and stress-induced immunosuppression (i.e., to the point where opportunistic pathogens and parasites overwhelm the host) must be met for animals to survive and become reproductively successful.

Several studies support this hypothetical construct. For instance, hamsters maintained in short photoperiods (and made to experience short-term stress) show more rapid wound healing and trafficking of leukocytes to the skin compared to unstressed short-day animals (or stressed and unstressed long-day hamsters). In one passerine bird, cutaneous immune responses to a T-cell mitogen (phytohemagglutinin, PHA) were less dramatic in animals induced into breeding than in those in a nonbreeding condition. Moreover, corticosterone appears to be immunosuppressive in the same species only in winter (not summer) conditions. Such immune–endocrine interrelationships are not always uniform between the sexes within a species. For example, acute stress enhances cell-mediated immune activity (delayed-type hypersensitivity to dinitrofluorobenzene) in female, but not male, Siberian hamsters. Indeed, ample evidence indicates that immune activity can vary dramatically between sexes; thus, it is plausible that the glucocorticoid mediators of immune defenses might be labile as well.

One other hormone that is integrally involved in vertebrate photoperiodism, melatonin, can affect the seasonal interplay between corticosterone and immune function. Generally, melatonin enhances immune activity, whereas glucocorticosteroids (at chronically high levels) compromise immune activity. Under certain circumstances, melatonin treatment may ameliorate the immunosuppressive effects of glucocorticosteroids. For instance, Siberian hamsters attenuated the duration of fever during simulated short winter day lengths, presumably to conserve energy. Hamsters housed in long days were injected with saline or melatonin before lights off each day for 1 or 6 weeks; then fever was assessed following injections of bacterial lipopolysaccharide (LPS). The fever duration was attenuated (32%) only in hamsters that decreased body mass, increased glucocorticoid concentrations, and displayed gonadal regression in response to 6 weeks of melatonin. These results suggest that long-term exposure to long-duration melatonin signals induce the physiological changes required for short-day immune responses, perhaps via interactions among hormones downstream from melatonin, such as cortisol and leptin.

In addition to their effects on immune function, corticosteroids can dramatically affect behavior, particularly behaviors associated with reproduction. Generally, stress responses are up- and down-regulated across seasons, but the way in which this happens depends on the social and physical environments in which species live. For instance, songbirds living at high latitudes are thought to mount weak corticosterone responses to stress because the abandonment of offspring that it might lead to would be maladaptive given that these species typically only have one opportunity per year to breed. In the tropics, however, corticosterone–behavior relationships may be different because reproductive efforts are generally less constrained by time.

Considered together, both the incidence and responses to stressors vary on a seasonal basis. The environmental regulation of seasonal changes in many traits, including reproduction, metabolism, and immunity, is primarily mediated by photoperiod. From a physiological prospective, the pineal hormone melatonin appears to coordinate seasonal adjustments in stress responses including changes in reproductive, metabolic, and immune function. To what extent the immune-enhancing effects of melatonin are unique to seasonally breeding rodents and whether it is a generalized feature of seasonal adjustments in stress responses among humans remain largely unspecified. Although physiological responses are important mediators of seasonal and photoperiodic changes in stress responses, behavioral alterations might also have an important role in reducing their energetic costs. The careful integration of the behavioral and physiological mechanisms underlying the seasonality of stress responsiveness at both ultimate and proximate levels should provide important and novel insights into the interaction among seasonal environmental factors, stressors, immune function, and the pattern of diseases and mortality. To date, most efforts have not advanced much beyond describing the seasonal changes that occur in responses to stressors. The next step is to identify the cellular and molecular mechanisms that mediate the effects of season and photoperiod on stress responsiveness and link variability in these mechanisms to fitness.

See Also the Following Articles

Alarm Phase and General Adaptation Syndrome; Allostasis and Allostatic Load; Glucocorticoids, Role in Stress; Immune Response; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress.

Further Reading

Bilbo, S. D., Dhabhar, F. S., Viswanathan, S. A., et al. (2002). Short day lengths augment stress-induced enhancement of skin immune function. *Proceedings of the National Academy of Sciences USA* **99**, 4067–4072.

Breuner, C. W. and Orchinik, M. (2002). Downstream from corticosterone: seasonality of binding globulins, receptors, and behavior in the avian stress response. In: Dawson, A. (ed.) *Avian endocrinology*, pp. 385–399. New Delhi: Narosa Publishing.

Breuner, C. W., Orchinik, M., Hahn, T. P., et al. (2003). Differential mechanisms for regulation of the stress response across latitudinal gradients. *American Journal of Physiology* **285**, R594–R600.

Martin, L. B., Gilliam, J., Han, P., et al. (2005). Corticosterone suppresses immune function in temperate but not tropical house sparrows (*Passer domesticus*). *General and Comparative Endocrinology* **140**, 126–135.

Neigh, G. N., Bowers, S. L., Pyter, L. M., et al. (2004). Pyruvate prevents restraint-induced immunosuppression via alterations in glucocorticoid responses. *Endocrinology* **145**, 4309–4319.

Nelson, R. J., Demas, G. E., Klein, S. L., et al. (2002). *Seasonal patterns of stress, immune function, and disease*. New York: Cambridge University Press.

Nelson, R. J. (2004). Seasonal immune function and disease responses. *Trends in Immunology* **24**, 187–192.

Niyogi, S., Biswas, S., Sarker, S., et al. (2001). Seasonal variation of antioxidant and biotransformation enzymes in barnacle, *Balanus balanoides*, and their relation with polyaromatic hydrocarbons. *Marine Environmental Research* **52**, 13–26.

Prendergast, B. J., Yellon, S. M., Tran, L. T., et al. (2001). Inhibition of cellular immune function by *in vitro* melatonin is modulated by photoperiodic history in Siberian hamsters. *Journal of Biological Rhythms* **16**, 224–233.

Romero, L. M. (2002). Seasonal changes in plasma glucocorticoid concentrations in free-living vertebrates. *General and Comparative Endocrinology* **128**, 1–24.

Romero, L. M. (2004). Physiological stress in ecology: lessons from biomedical research. *Trends in Ecology and Evolution* **19**, 249–255.

Sapolsky, R. M., Romero, L. M. and Munck, A. (2000). How do glucocorticoids influence stress responses?: integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews* **21**, 55–89.

Yellon, S. M., Kim, K., Hadley, A. R., et al. (2005). Time course and role of the pineal gland in photoperiod control of innate immune cell functions in male Siberian hamsters. *Journal of Neuroimmunology* **161**, 137–144.

Seasonal Rhythms

L M Romero
Tufts University, Medford, MA, USA

© 2007 Elsevier Inc. All rights reserved.

Evidence for Seasonal Glucocorticoid Rhythms
Mechanisms Regulating Annual Rhythms of Glucocorticoids
Why Do Seasonal Glucocorticoid Rhythms Exist?
Adaptive Significance

Glossary

Basal glucocorticoids	Plasma concentrations of glucocorticoids in the absence of a stressor.
Corticotropin-releasing hormone (CRH)	Hormone released by the hypothalamus.
Glucocorticoid-binding globulin (CBG)	Proteins in the plasma that bind glucocorticoids.
Hypothalamic-pituitary-adrenal (HPA) axis	The neuroendocrine pathway leading from the brain to the adrenals that culminates in glucocorticoid release.
Permissive effect	An indirect effect whereby a hormone indirectly allows another system to work better.
Stressor	A stimulus that elicits a release of glucocorticoids.

Evidence for Seasonal Glucocorticoid Rhythms

Glucocorticoid secretion is often considered an obligatory response to stressors. In a laboratory setting, stressors applied in similar circadian, developmental, and psychological contexts within the same species produce remarkably consistent magnitudes of glucocorticoid secretion. Recent evidence, however, indicates that glucocorticoid release is modulated seasonally in the majority of free-living species that have been studied to date. In other words, the magnitudes of both basal and stress-induced glucocorticoid concentrations vary throughout the year. An example, data for white-crowned sparrows (*Zonotrichia leucophrys*), is presented in Figure 1.

Many recent studies have measured glucocorticoids in free-living wild animals by directly sampling from freshly caught individuals. Plasma samples collected within a few minutes of capture are assumed to reflect prestress glucocorticoid concentrations because glucocorticoids start to increase 2–3 min after initiation of a stressor. Stress-induced glucocorticoid concentrations then reflect the animal's response to capture, handling, captivity, etc. and are assumed to be a measure of the hypothalamic-pituitary-adrenal (HPA) axis's ability to respond to a stressor.

These studies provide clear evidence that a majority of reptile, amphibian, avian, and mammal species seasonally modulate glucocorticoid concentrations. Males and females sometimes differ in whether they show seasonal rhythms, and occasionally a species will show a seasonal rhythm in basal but not stress-induced glucocorticoid concentrations (or vice versa). However, approximately 90% of all species that have been studied have at least one sex with a seasonal rhythm of glucocorticoid release.

The evidence for an annual glucocorticoid rhythm is generally much stronger for basal concentrations, primarily because more studies have focused on this time point. Furthermore, avian species often show a seasonal rhythm in basal but not stress-induced glucocorticoid concentrations, suggesting that basal and stress-induced glucocorticoid concentrations are regulated differently. Laboratory studies show that basal and stress-induced glucocorticoids have different regulatory mechanisms and serve different physiological functions that are mediated by separate receptors. It is not surprising, therefore, that the HPA axis might be regulated differently under basal and stress conditions, but these regulatory pathways can be further dissociated into different annual cycles.

Of those species that do show a seasonal rhythm, the seasonal peak is usually during the breeding period. There are three notable exceptions. First, the most robust rhythm in birds is a nadir during the prebasic molt when they replace their feathers. This is thought to protect the bird from the protein catabolism effects of glucocorticoids during the protein mobilization of feather replacement. Second, three desert-breeding birds show a seasonal glucocorticoid peak during the rainy season rather than the breeding period, suggesting that seasonal rhythms can be timed to environmental cues as well as seasonal physiological cues. Third, even though most mammal species show a seasonal rhythm, there is no consensus season when mammals tend to have elevated glucocorticoid concentrations.

The capacity of glucocorticoid-binding globulins (CBGs) has also been found to vary seasonally and

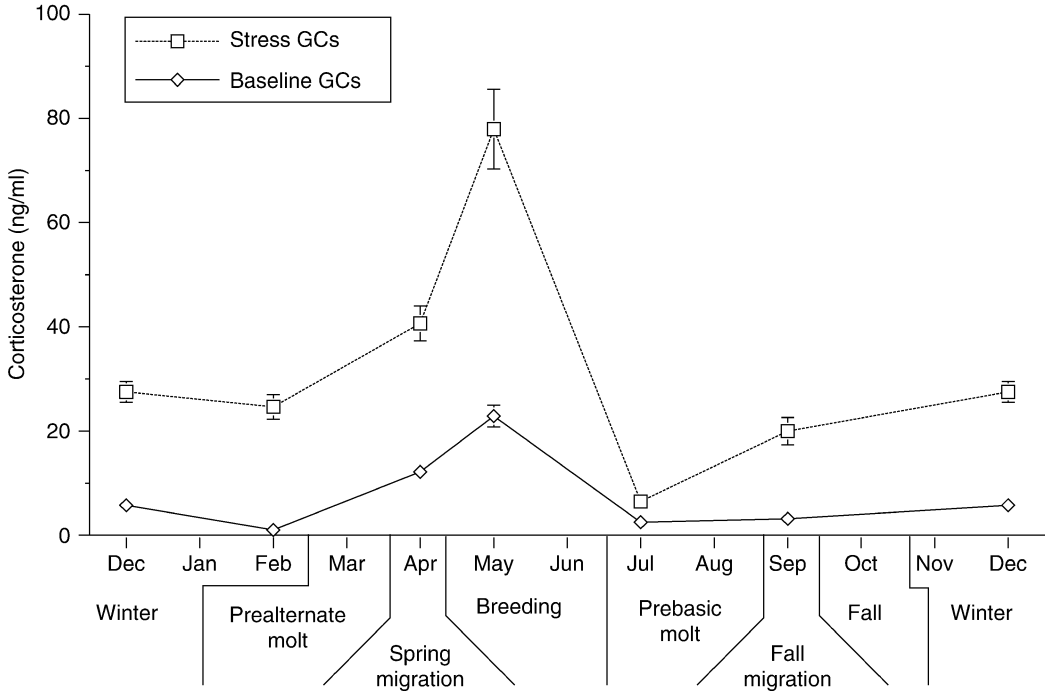


Figure 1 Seasonal rhythm of baseline and stress-induced corticosterone from white-crowned sparrows. Reprinted from Romero, L. M. (2002). Seasonal changes in plasma glucocorticoid concentrations in free-living vertebrates. *General and Comparative Endocrinology* **128,** 1–24 with permission from Elsevier.

to be regulated by gonadal testosterone. Because only free steroid is generally considered to be available for diffusion into tissues, free steroid concentrations may be more biologically relevant than total steroid concentrations. When total and free steroid concentrations are compared, the seasonal rhythm in total glucocorticoids often changes or disappears entirely.

Many wild species held in long-term captivity also show seasonal rhythms, but these rhythms do not always match what is found in free-living animals. This likely reflects that both domestication and captivity alter HPA axis function in ways that are different for each species, making it difficult, if not impossible, to ascertain natural annual variation from these populations.

Mechanisms Regulating Annual Rhythms of Glucocorticoids

Progress is just beginning in understanding how HPA axis function is regulated seasonally. What induces these changes, be it photoperiod, temperature, food availability, etc., is currently unknown. Plasma glucocorticoid concentrations often are correlated with adrenal mass, and early studies showed increases in adrenal mass during breeding. Consequently, the ability of adrenal tissue to respond to adrenocorticotropic hormone (ACTH) changes seasonally. A few studies in birds indicate that there are multiple regulatory points in the HPA axis. Some species seasonally regulate glucocorticoid release from the adrenal, others seasonally regulate ACTH release from the pituitary, and still others regulate CRH and arginine vasotocin (the avian congener of arginine vasopressin) release from the hypothalamus. These regulatory control points may be related to how sensitive each species is to adverse environmental conditions.

Glucocorticoids also interact with the gonadal system. Although glucocorticoids inhibit gonadal hormone release, exogenous testosterone elevates glucocorticoids in free-living birds. This suggests a complex interaction between the gonadal and adrenal systems, with glucocorticoids downregulating gonadal androgen release concurrent with gonadal androgens elevating glucocorticoid release. Gonadal androgens, therefore, could potentially be an important physiological regulator of seasonal glucocorticoid rhythms.

Why Do Seasonal Glucocorticoid Rhythms Exist?

Modulation of glucocorticoid release presents an interesting paradox: if glucocorticoid release is so important for survival, then how do animals survive stressors at certain times of the year when they essentially fail to release glucocorticoids? There have been three major hypotheses proposed to explain seasonal glucocorticoid rhythms.

Energy Mobilization Hypothesis

The energy mobilization hypothesis focuses on the metabolic effects of glucocorticoids. Glucocorticoids play an important role in energy mobilization, especially during stress, so glucocorticoid concentrations should be highest during energetically costly times of the year. According to this hypothesis, high glucocorticoid concentrations during breeding result from breeding exerting the highest energetic demands during the year. Females of many species, for example, expend substantial energy during breeding, and breeding males of many species have increased energy costs associated with testosterone and territorial defense. Moreover, increased energy mobilization could explain seasonal peaks during nonbreeding periods if these periods are more energetically costly than breeding.

Two examples, however, mammalian hibernation and avian molt, highlight the insufficiency of the energy mobilization hypothesis in providing a universal explanation for seasonal glucocorticoid rhythms. Hibernating mammals require dramatically increased fat stores to survive the winter, but it is not clear that these energy requirements are greater or less than lactation. Second, feather replacement during molt in birds requires substantial energy expenditure, yet this period is the nadir of the annual rhythm. The energy mobilization hypothesis can be a powerful explanatory mechanism for certain species during some times of the year, but it clearly is insufficient to fully explain seasonal glucocorticoid rhythms.

Behavior Hypothesis

The behavior hypothesis focuses on the behavioral effects of glucocorticoids. It posits that annual glucocorticoid rhythms result from different requirements for expressing (or not expressing) glucocorticoid-mediated behaviors at different times of the year. For example, an important glucocorticoid-mediated behavior in wild animals might be fleeing an area and relocating during storms. This may be an excellent strategy for most of the year, but could be catastrophic for an individual's overall fitness when relocation requires abandoning young. Consequently, the behavior hypothesis suggests that the need to seasonally regulate the expression of those behaviors drives glucocorticoid rhythms.

The behavior hypothesis, however, does not fit a more general annual rhythm in glucocorticoid concentrations. Glucocorticoid concentrations are highest in most species during breeding, exactly when glucocorticoid concentrations should be lowest if seasonal glucocorticoid rhythms were primarily to regulate the behavioral consequences of glucocorticoid release. Consequently, this hypothesis may be useful in explaining modulation of glucocorticoid concentrations within a single season, but not across seasons.

Preparative Hypothesis

The preparative hypothesis focuses on the coordinating aspects of glucocorticoid physiology and posits that seasonal rhythms in glucocorticoid concentrations serve to modulate the priming of other stress pathways during periods with different potential exposure to stressors. Glucocorticoids are only one part of the stress response. Other hormones (especially epinephrine and norepinephrine), neurotransmitters (e.g., 5-HT, CRH), opioid peptides, and cytokines (e.g., IL-6), as well as other brain functions, exert their effects rapidly with the onset of stress without involving the pituitary or adrenals. However, glucocorticoids have a permissive effect on many of these systems and thereby facilitate better performance under stress. Higher basal glucocorticoid concentrations may augment the priming effects on these systems in preparation for further stressors. This is the traditional explanation for timing the circadian peak for the beginning of the active period. These permissive effects may help prepare the organism for subsequent stress responses.

If the preparative hypothesis is correct, higher glucocorticoid concentrations should coincide with seasons in which predictable stressors are more likely. The available data fit this prediction moderately well. The risk of exposure to many stressors in wild animals likely varies seasonally. During the breeding season, for example, predation risk may increase when caring for young; competition for mates, with the risk of frequent fights, increases; and disease may become more prevalent when breeding individuals congregate. Consequently, glucocorticoid concentrations may vary seasonally in order to mediate changes in preparedness in the nonglucocorticoid pathways of the stress response. However, the existence of predictable seasonal increases in stressor risk has not yet been demonstrated.

Adaptive Significance

Proper concentrations of glucocorticoids seem crucial for survival. Too little can result in even mild stressors killing animals, but too much can lead to suppressed gonadal and immune function and ultimately disease. Consequently, a balance must be struck between needing glucocorticoids to survive stressors and modulating glucocorticoid secretion to prevent deleterious exposures. What is becoming clear is that there is a seasonal rhythm in what is the

right amount of glucocorticoids. Understanding why the right amount changes seasonally will help us understand why glucocorticoids are released during stress and how they help animals survive. The three explanatory hypotheses presented earlier are derived from different aspects of glucocorticoid physiology. The energy mobilization hypothesis emphasizes glucocorticoid's metabolic effects; the behavior hypothesis emphasizes the desired acute behavioral effects of glucocorticoids; and the preparative hypothesis emphasizes the regulatory role that glucocorticoids have in other stress systems. Integrating these hypotheses into a cohesive physiology will provide the foundation for understanding how elevated glucocorticoid concentrations aid in survival.

See Also the Following Articles

Circadian Rhythms, Effects of Prenatal Stress in Rodents; Circadian Rhythms, Genetics of; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress; Corticosteroids and Stress; Cortisol Awakening Response; Corticosteroid Receptors; Glucocorticoid Negative Feedback; Hypothalamic-Pituitary-Adrenal.

Further Reading

Guillette, L. J., Cree, A. and Rooney, A. A. (1995). Biology of stress: interactions with reproduction, immunology and intermediary metabolism. In: Warwick, C., Frye, F. & Murphy, J. B. (eds.) *Health and welfare of captive reptiles*, pp. 32–81. London: Chapman & Hall.

Romero, L. M. (2002). Seasonal changes in plasma glucocorticoid concentrations in free-living vertebrates. *General and Comparative Endocrinology* **128**, 1–24.

Sapolsky, R. M., Romero, L. M. and Munck, A. U. (2000). How do glucocorticoids influence stress-responses? Integrating permissive, suppressive, stimulatory, and adaptive actions. *Endocrine Reviews* **21**, 55–89.

Tyrrell, C. L. and Cree, A. (1998). Relationships between corticosterone concentration and season, time of day and confinement in a wild reptile (tuatara, *Sphenodon punctatus*). *General and Comparative Endocrinology* **110**, 97–108.

Wingfield, J. C. (1994). Modulation of the adrenocortical response to stress in birds. In: Davey, K. G., Peter, R. E. & Tobe, S. S. (eds.) *Perspectives in comparative endocrinology*, pp. 520–528. Ottawa: National Research Council of Canada.

Wingfield, J. C. and Romero, L. M. (2001). Adrenocortical responses to stress and their modulation in free-living vertebrates. In: McEwen, B. S. & Goodman, H. M. (eds.) *Handbook of physiology; section 7: The endocrine system; vol. IV: Coping with the environment: neural and endocrine mechanisms*, pp. 211–234. New York: Oxford University Press.

Wingfield, J. C., Hunt, K., Breuner, C., et al. (1997). Environmental stress, field endocrinology, and conservation biology. In: Clemmons, J. R. & Buchholz, R. (eds.) *Behavioral approaches to conservation in the wild*, pp. 95–131. Cambridge, UK: Cambridge University Press.

Secretagogue

G Fink
Mental Health Research Institute, Melbourne, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Fink, volume 3, pp 409–411, © 2000, Elsevier Inc.

Principles
Secretagogues Involved in the Neuroendocrine Response to Stress
Multiple Secretagogues in Other Systems: Exemplified by Growth Hormone Secretion
Sympathetic System: Role in Fight or Flight and Model for Exocytosis
Receptors and Intracellular Signaling Cascades
Conclusion

Glossary

Catecholamines — Adrenaline and noradrenaline are catecholamines. Neurotransmitters of the sympathetic nervous system (noradrenaline) and the adrenal medulla (adrenaline), they rapidly prepare the body for flight or fight in response to a stress. This is achieved mainly by adrenaline or noradrenaline stimulation of glucose release from the liver, increased heart rate, and redistribution of the blood circulation (by tissue-selective vasomotor control) from nonessential (e.g., the gut) to essential organs such as the brain, heart, and skeletal muscles.

Chemical neurotransmitter — A compound synthesized and released by a nerve cell (neuron) that activates or inhibits another neuron or other type of excitable target cell such as muscle.

Chromagranins	A class of acidic proteins present in large secretory granules (vesicles) in neuroendocrine tissues and neuroendocrine-derived tumors. Chromagranins were first identified in the chromaffin cells of the adrenal medulla, sympathetic nerves, and the pituitary gland. One of the chromagranins, secretogranin II, is the precursor of secretoneurin, a neuropeptide with a wide distribution in brain through all vertebrate phyla.
Endocrine glands	Glands that secrete signaling molecules, termed hormones, into the bloodstream, which transports the hormones to their target organs and tissues, where they exert their effects.
Exocrine glands	Glands that secrete to the outside of the body (i.e., into the gut, nasopharynx, bronchial tree, genitourinary system, and skin) as opposed to endocrine glands, which secrete hormones into the bloodstream. The acinar cells of the pancreas, which secrete gut enzymes, are key examples of exocrine cells in the context of this topic.
Exocytosis	The release of a secretory product, such as neurotransmitters, in packets or quanta. The secretory products are packaged by the Golgi apparatus of the cell in membrane-bound vesicles; exocytosis involves fusion of the vesicular membrane with the cell membrane, resulting in the release of the quantum of product inside the vesicle. This is thought to be the main mechanism of secretion in most secretory cells, including neurons.
Ligand	A molecule that binds or bonds specifically to another substance/compound – in biology, usually a receptor.
Neurohormone	A compound synthesized and released by a nerve cell that reaches its target by transport in special vessels, such as the hypothalamo-hypophysial vessels in the case of the hypothalamic-pituitary system, or the systemic circulation for peptides such as oxytocin and vasopressin whose targets are in the periphery. Most, but not all, neurohormones are peptides.
SNAP-25	Synaptosome-associated protein of 25 kDa; one of the SNARE proteins involved in calcium-regulated exocytosis.
SNARE proteins	Soluble N-ethylmaleimide-sensitive factor attachment protein receptors; involved in calcium regulated exocytosis.
Stimulus secretion coupling	The link between a stimulus, usually membrane depolarization (axon potentials in the case of nerve cells) under physiological conditions, and the release of a neurotransmitter or neurohormone.

Principles

A secretagogue is an agent that promotes the secretion of hormones, neurohormones, chemical neurotransmitters, enzymes, or other compounds synthesized and secreted by cells. The selectivity or specificity of a secretagogue is determined by the structure of the secretagogue and that of its receptors on or within the cell. Secretagogues may be endogenous biological agents natural to the animal, exogenous compounds obtained from plants, or artificially synthesized pharmacological agents. The activity of exogenous secretagogues is thought to be due to the fact that the agent, by virtue of its chemistry and conformation, is capable of activating receptors specific to an endogenous secretagogue. There are numerous examples of this type of pharmacological mimicry: for example, acetylcholine (endogenous) and nicotine or muscarine (plant derived), opiates (plant derived), and the endogenous opioids – enkephalins and endorphins. The different classes of receptors that mediate secretagogue action are outlined later in the article.

Some secretagogues are nonspecific and will stimulate the release of most secretory substances. The most widely known and commonly used (experimentally) nonspecific secretagogue is a high concentration of extracellular potassium ions. High extracellular potassium concentrations trigger secretion by depolarizing the cell membrane, which induces a massive increase in intracellular calcium ions. Studies with calcium ionophores (compounds that induce calcium influx into cells) and calcium channel modulators in adrenaline-secreting chromaffin cells suggest that it is the rate of calcium ion influx rather than the absolute level of intracellular calcium ion reached that determines the amount of catecholamine released. Secretagogue-induced depolarization of cell membranes is associated with the release of neurotransmitters, neurohormones, some pancreatic and salivary gland enzymes, and the pancreatic hormone insulin. The coupling of stimulus-induced membrane depolarization with secretion is termed stimulus–secretion coupling, a term coined by W. W. Douglas based on his work on the neurohypophysis and the adrenal medulla.

Secretagogues Involved in the Neuroendocrine Response to Stress

Several secretagogues are involved in organizing the biological response to stress. Thus, stress, most frequently generated by an event recognized by the senses as dangerous or potentially dangerous, activates neuronal systems in higher brain centers that activate the main components of the limbic system, which include the hippocampus, amygdala, and

hypothalamus. This activation and the transmission of signals between nerve cells (neurons) involve the release of chemical neurotransmitters at nerve junctions (synapses). Nerve impulses generated as a consequence of stress activation of higher brain centers and the limbic system are transmitted to the paraventricular nuclei (PVN) of the hypothalamus, the major stress nuclei of the brain. These nuclei comprise neurons that synthesize and secrete corticotropin-releasing hormone-41 (CRH-41) and arginine vasopressin (AVP) and project to the median eminence of the hypothalamus, where they terminate on the primary plexus of the hypophysial portal vessels. Released from PVN terminals in the median eminence and transported to the anterior pituitary gland by the hypophysial portal vessels, CRH-41 and AVP act synergistically to stimulate the release of adrenocorticotropin (ACTH) into the systemic circulation. The PVN neurons, therefore, constitute the final common pathway neurons of the central neuroendocrine stress control system. Final common pathway neurons was a term first used by Sherrington to describe the alpha motor neurons of the spinal cord that project to and control skeletal muscle activity.

In humans and rodents, CRH-41 is the main secretagogue for ACTH release; its action is potentiated by AVP. In sheep, the situation is reversed, with AVP being the main ACTH secretagogue and CRH-41 potentiating the action of AVP. In addition to these ACTH-releasing factors, there is experimental evidence that under certain conditions the brain may inhibit ACTH release. A likely candidate for an ACTH-inhibiting factor is the cardiac hormone atrial natriuretic peptide (ANP), which is also synthesized in PVN neurons and released into hypophysial portal blood.

Transported by the circulation to the adrenal glands, ACTH is the main secretagogue for stimulating the release of adrenal corticosteroids, and especially the glucocorticoids, cortisol in humans and corticosterone in rodents. Under certain situations, steroids can be secretagogues. Thus, estrogen, in its positive feedback mode for gonadotropin release, induces the release of gonadotropin-releasing hormone and prolactin, albeit possibly by indirect actions. Estrogen can also potentiate compound 48/80- or substance P-induced release of serotonin and histamine from mast cells. Glucocorticoids are potent inhibitors of ACTH release; this action is exerted directly on the anterior pituitary gland, where glucocorticoids inhibit the action of CRH-41 and AVP, as well as on the hypothalamus, where they inhibit CRH-41 and AVP release. This inhibitory effect of glucocorticoids is the basis for their negative feedback control of ACTH release. Whether corticosteroids are secretagogues is not known; however, by exerting pronounced effects on protein synthesis and enzyme activity in nerve and other cells, corticosteroids modulate the action of secretagogues.

In addition to what might be called the classical secretagogues of the brain-pituitary-adrenal system, other secretagogues implicated in the control of stress-induced glucocorticoid secretion include urocortin (a homolog of CRH-41), angiotensin II, and cytokines such as tumor necrosis factor-α (TNF-α) and interleukin-1β (IL-1β). TNF-α and IL-1β both stimulate ACTH release and are two of several secretagogues that link the immune and neuroendocrine defense systems. Extreme stress, such as that produced by multiple trauma or septic shock, results in a paradoxical decrease of the secretion of ACTH in the presence of increased cortisol secretion. It is postulated that this paradox may be brought about by the inhibition of ACTH secretion by ANP and by the direct adrenocortical stimulation of glucocorticoid secretion by the peptide endothelin-1.

Multiple Secretagogues in Other Systems: Exemplified by Growth Hormone Secretion

The existence of more than one secretagogue for pituitary hormones is not confined to ACTH. Thus, growth hormone (GH) release is regulated by two hypothalamic peptides or neurohormones: GH-releasing hormone, which stimulates GH release, and somatostatin, which inhibits GH release. Ghrelin, a peptide secreted by the stomach, is also an endogenous ligand for the GH secretagogue receptor. Ghrelin appears to be present also in the hypothalamus, but its precise mechanism of action and interactions with GH-releasing hormone and somatostatin needs to be elucidated. Finally, GH release is also induced by increased plasma concentrations of certain amino acids, such as arginine and lysine (which can therefore be regarded as GH secretagogues). However, while arginine (administered intravenously) is used as a provocative test of GH secretion in human subjects, there appears to be no robust evidence that amino acids administered orally stimulate detectable sustained GH release.

Sympathetic System: Role in Fight or Flight and Model for Exocytosis

The sympathetic component of the autonomic nervous system and the adrenal medulla play an equally important role in the stress response, especially in the acute phase when autonomic activation prepares the animal for fight or flight. That is, in response to stress, the brain through the hypothalamus and hindbrain

centers activates the sympathetic nervous system, which, through the release of noradrenaline at sympathetic nerve terminals and adrenaline from the adrenal medulla, stimulates several physiological changes that include an increase in heart rate and blood pressure and the rapid mobilization of glucose production by the liver.

Stimulation of adrenaline secretion from the adrenal medulla by the preganglionic secretagogue acetylcholine has been used extensively as a model for the investigation of secretagogue action and stimulus–secretion coupling. Based on the seminal electrophysiological studies of Bernard Katz and associates and confirmed by ultrastructural (electron microscopical) findings, it has long been accepted that stimulus–secretion coupling results in the release of neurotransmitters, hormones, and enzymes by way of exocytosis. Work on the adrenal medulla and sympathetic nerves proved important for elucidating the mechanism of exocytosis and the recycling and reuse of vesicular membranes, the characterization of the chromagranins, and the identification of peptides involved in exocytosis such as syntaxin, SNAP-25 and L- and N-type calcium channels. In brief, secretory vesicles formed at the trans-Golgi network of neuroendocrine and endocrine cells undergo translocation, docking, and priming before they are ready to fuse with the plasma membrane and release their secretory product into the extracellular space. This process, termed regulated exocytosis, is controlled by Ca^{2+} (using synaptotagmin) and is mediated by SNARE proteins. For details about this and vesicle priming, depriming and cell-membrane fusion see Stojilkovic (2005).

Receptors and Intracellular Signaling Cascades

There are two broad classes of receptors – cell surface and nuclear. Located intracellularly, the nuclear receptor superfamily binds steroid (e.g., cortisol, corticosterone, aldosterone, estradiol, testosterone) and thyroid hormones, derivatives of vitamins A and D, and xenobiotics. Cell surface receptors are composed of the following classes (examples of ligands):

1. Ligand-gated ion channels (e.g., nicotinic acetylcholine receptor)
2. Receptor tyrosine kinases (e.g., insulin, insulin-like growth factor, epidermal growth factor)
3. Receptor serine/threonine kinases (e.g., activin, inhibins)
4. Receptor guanylate cyclase (e.g., atrial natriuretic peptide)
5. G-protein-coupled receptor superfamily (e.g., adrenergic agents, muscarinic acetylcholine agents, peptide hormones and neurohormones, glycoproteins, glucagon, parathyroid hormone)
6. Cytokine receptors (e.g., growth hormone, prolactin, leptin)

With the exception of glucocorticoids, cell surface receptors are the targets of all the secretagogues involved in the neuroendocrine response to stress. Thus, CRH-41, vasopressin, ACTH, adrenaline, and noradrenaline all bind to specific G protein-coupled receptors, while ANP binds to guanylate cyclase receptors. Acetylcholine, the preganglionic secretagogue that triggers adrenaline release from the adrenal medulla, acts by way of nicotinic acetylcholine receptors on adrenal medullary cells.

Secretion induced by ligand activation of non-voltage-dependent cell surface receptors is triggered by one of a series of intracellular signaling cascades, the most well known of which are the adenylyl cyclase, phosphatidyl inositol, and kinase systems. As indicated previously, secretagogue-induced secretion involves a rapid increase in the intracellular concentrations of calcium ions, triggered either by effects on the cell surface calcium channels in the case of voltage-dependent receptors, or by the intracellular signaling cascades in the case of non-voltage-dependent receptors. In addition to stimulating secretion, activation of cell surface receptors frequently results in gene transcription that can lead to synthesis of secretory product or a key enzyme involved in the synthesis of secretory product. The latter is exemplified by the fact that acetylcholine, which induces the secretion of adrenaline from the adrenal medulla and noradrenaline from postganglionic neurons of the sympathetic system, also induces the synthesis of tyrosine hydroxylase, the rate-limiting enzyme for the synthesis of noradrenaline and adrenaline. This action of acetylcholine appears to be mediated by both nicotinic voltage-dependent receptors and muscarinic G protein-coupled receptors.

Conclusion

The mechanism of action of secretagogues is the subject of intensive research and is one of the major themes of modern molecular biology and molecular pharmacology. Calcium ions play a pivotal role in the action of secretagogue-induced secretion by exocytosis. Although a good deal is known about secretagogue action, a more rigorously defined understanding is likely to emerge from modern molecular genetic and microphysiological approaches.

See Also the Following Articles

Adrenal Medulla; Feedback Systems; Glucocorticoid Negative Feedback; Glucocorticoids, Role in Stress; Neuroendocrine Systems.

Further Reading

Craske, M., Takeo, T., Gerasimenko, O., et al. (1999). Hormone-induced secretory and nuclear translocation of calmodulin: oscillations of calmodulin concentration with the nucleus as an integrator. *Proceedings of the National Academy of Sciences USA* **96**, 4426–4431.

Douglas, W. W. (1968). Stimulus secretion coupling: the concept and clues from chromaffin and other cells. *British Journal of Pharmacology* **34**, 453–474.

Douglas, W. W. (1973). How do neurones secrete peptides? Excocytosis and its consequences, including "synaptic vesicle" formation, in the hypothalamo-neurohypophysial system. *Progress in Brain Research* **39**, 21–39.

Douglas, W. W. (1974). Involvement of calcium in excocytosis and the excocytosis-vesiculation sequence. *Biochemical Society Symposia* **39**, 1–28.

Eder, U., Fischer-Colbrie, R., Kogner, P., et al. (1998). Levels and molecular forms of chromagranins in human childhood neuroblastomas and ganglioneuromas. *Neuroscience Letters* **253**, 17–20.

Fink, G., Robinson, I. C. A. F. and Tannahill, L. A. (1988). Effects of adrenalectomy and glucocorticoids on the peptides CRF-41, AVP and oxytocin in rat hypophysial portal blood. *Journal of Physiology* **401**, 329–345.

Fink, G., Dow, R. C., Casley, D., et al. (1992). Atrial natriuretic peptide is involved in the ACTH response to stress and glucocorticoid negative feedback in the rat. *Journal of Endocrinology* **135**, 37–43.

Gerasimenko, O. V., Gerasimenko, J. V., Petersen, O. H., et al. (1996). Short pulses of acetylcholine stimulation induce cytosolic Ca^{2+} signals that are excluded from the nuclear region in pancreatic acinar cells. *Pflugers Archive* **432**, 1055–1061.

Larsen, P. R., Kronenberg, H. M., Melmed, S. and Polonsky, K. S. (eds.) (2003). *Williams textbook of endocrinology* (10th edn.) Philadelphia, PA: Saunders.

Leitner, B., Schneitler, C., Klocker, H., et al. (1998). Formation and sequence analysis of secretoneurin, a neuropeptide derived from secretogranin 11, in mammalian, bird, reptile, amphibian and fish brains. *Neuroscience Letters* **248**, 105–108.

Leitner, B., Lovisetti-Scarnihorn, P., Heilmann, J., et al. (1999). Subcellular localization of chromagranins, calcium channels, amine carriers, and proteins of the exocytotic machinery in bovine splenic nerve. *Journal of Neurochemistry* **72**, 1110–1116.

Mogami, H., Nakano, K., Tepikin, A. V., et al. (1997). Ca^{2+} flow via tunnels in polarized cells: recharging of apical Ca^{2+} entry through basal membrane patch. *Cell* **88**, 49–55.

Petersen, O. H. (1996). Can Ca^{2+} be released from secretory granules on synaptic vesicles? *Trends in Neuroscience* **19**, 411–413.

Petersen, O. H. and Iwatsuki, N. (1978). The role of calcium in pancreatic acinar cell stimulus-secretion coupling: an electrophysiological approach. *Annals of the New York Academy of Sciences* **307**, 599–617.

Petersen, O. H., Petersen, C. C. and Kasai, H. (1994). Calcium and hormone action. *Annual Review of Physiology* **56**, 297–319.

Stojilkovic, S. S. (2005). Ca(2+)-regulated exocytosis and SNARE function. *Trends in Endocrinology and Metabolism* **16**, 81–83.

Turnbull, A. V., Dow, R. C., Hopkins, S. J., et al. (1994). Mechanisms of activation of the pituitary-adrenal axis by tissue injury in the rat. *Psychoneuroendocrinology* **19**, 165–178.

Vliagotis, H., Dimitriadou, V., Boucher, W., et al. (1992). Estradiol augments while tamoxifen inhibits rat mast cell secretion. *International Archives of Allergy and Immunology* **98**, 398–409.

Seizures *See:* Epilepsy.

Selective Serotonin Reuptake Inhibitors (SSRIs)

M Gitlin
Geffen School of Medicine at the University of California, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Serotonin and the Selective Serotonin Reuptake Inhibitors
Stress and the Brain
Stress and Serotonin
Antidepressants, SSRIs, and Stress Responses

Glossary

Brain-derived neurotrophic factor (BDNF)	A brain peptide that enhances growth of neurons.
Hippocampus	A brain structure vital for learning and memory.
Selective serotonin reuptake inhibitors (SSRIs)	A class of antidepressants that have the dominant initial biological effect of increasing serotonin in the synaptic cleft.
Serotonin	A neurotransmitter that is involved in the regulation of mood, anxiety, and stress responsiveness, among other effects.

Serotonin and the Selective Serotonin Reuptake Inhibitors

Selective serotonin reuptake inhibitors (SSRIs) are the dominant class of antidepressants prescribed in the Western world. Although they are not more effective than other classes of antidepressants, their ease of use (most patients show optimal benefit from one daily dose of any of the SSRIs), their relatively benign side effect profile, and the extraordinary marketing campaigns by the pharmaceutical manufacturers of the individual agents have made them the most commonly prescribed antidepressants and among the most commonly prescribed agents in all of medicine.

All SSRIs share the common trait of blocking the serotonin transporter, the receptor on the presynaptic neuron that regulates the reuptake of serotonin back into the presynaptic neuron. The functional consequence of blocking serotonin reuptake is to allow serotonin to remain in the synaptic cleft longer, thereby enhancing its likelihood of hitting the postsynaptic serotonin receptor. The end result of this is to enhance serotonin neurotransmission. The strength of serotonin activity at serotonin receptors is inversely proportional to the number of serotonin transporter molecules present at the presynaptic membrane. Although the mechanism by which SSRIs ameliorate depression is far more complex than this, the initial enhancement of serotonergic function by reuptake blockade is generally thought to be the first step in a cascade of biological effects.

Although originally released and marketed as treatments for depression, SSRIs have now been demonstrated to be effective in a wide variety of psychiatric disorders, including panic disorder, social anxiety disorder, obsessive-compulsive disorder, generalized anxiety disorder, posttraumatic stress disorder (PTSD), and some of the symptoms of certain personality disorders. Thus, these powerful serotonin enhancers seem to be effective treatments in a wide variety of mood and anxiety disorders. Many, but not all, of these disorders can be triggered or exacerbated by acute or chronic stress. For some individuals – specifically those whose depressive episodes are triggered by acute or chronic stressful life events – depression can be described as the end product of failed adaptation to chronic emotional stress. Because of the overlap among anxiety disorders, depression, and stress, the remainder of this article reviews all three, even though, of course, these are similar but not identical constructs or clinical conditions.

Stress and the Brain

Stress alters the brain in a variety of ways. Two of these stress-related effects – decreases in hippocampal volume and the diminishing of the activity of a brain peptide, brain-derived neurotrophic factor (BDNF) – are highly relevant in understanding the potential effects of antidepressants and SSRIs, in particular in their altering responses to stress.

A decrease in the volume of the hippocampus, the brain structure vital for learning and memory, is among the most consistent consequences of stress. In rodents, experimental chronic stress causes the loss of hippocampal neurons. Similar findings have been demonstrated in humans suffering from depression and in other stress-related psychiatric disorders such as PTSD. Lower hippocampal volumes in depressed patients compared with normal people are seen bilaterally, although this finding is more consistent on the left versus right hippocampus. Similar findings are seen with PTSD, even in those patients not suffering from comorbid depression. In addition, in two studies, the reduction in hippocampal volume

correlated with the length of the depressive illness. It is as yet unclear whether these changes in actual brain structure seen with stress versus psychiatric disorders such as depression or PTSDs are similar or identical.

The second important stress-related effect on brain function is in the decreasing activity of the brain peptide BDNF. BDNF plays a major role in the survival, growth, and protection of neurons, both in the central and the peripheral nervous systems. The role of BDNF in neuronal function continues into adult life; thus, BDNF should be considered as playing an ongoing regulatory role for optimal neuronal function in youngsters and adults. Normally, neurons are rewired and shaped (with the sprouting of dendrites, called dendritic arborization) throughout adult life. BDNF seems integrally involved with this process. In addition, the production of new neurons in adults, called neurogenesis, seems also to be regulated in part by BDNF. Of note, BDNF shows the greatest expression in the hippocampus, the brain region most consistently demonstrated to show morphological changes in stress and depression.

In animals, a variety of stressors (such as forced immobilization in rats or learned helplessness following exposure to inescapable shock) consistently cause a dramatic reduction in BDNF expression in the hippocampus. There is some evidence that the prolonged glucocorticoid secretion (see **Glucocorticoids, Effects of Stress on, Glucocorticoids, Role in Stress**) that is consistently observed in association with severe stress may mediate some of the effects of stress and depression on decreased BDNF function.

Human studies of BDNF are more difficult to obtain because direct measurements of brain are not possible to obtain and the relationship between blood and brain levels of BDNF is still uncertain. Nonetheless, in one study, serum BDNF levels in depressed patients were reported to be lower than those seen in normal controls. In addition, in another study, serum BDNF levels correlated inversely with depression rating scale scores.

Stress and Serotonin

Central nervous system responses to stress are mediated by multiple neurotransmitter systems (see **Primates: Rearing and Effects of Stress on Primate CNS Function, Stress and CNS Arousal; Genomic Contributions**). Consistent evidence suggests that the serotonin system plays an important role in the brain's response to stress, thereby further suggesting that modifying serotonergic function may alter responses to stress. In animals, low serotonin is associated with increased sensitivity to environmental cues and greater responsiveness to threat. In humans, this implies that individuals with low serotonin may be more easily traumatized by stressful life events (or have a lower threshold for showing the effects of traumatic events) than those with higher baseline serotonin levels.

The likelihood of stress triggering depressive symptoms and disorders seems to be moderated in part by different profiles of the gene that helps regulate the serotonin transporter (i.e., the receptor that regulates the reuptake of serotonin into the presynaptic neuron). In a seminal study, young adults with either one or two copies of the short allele of the gene that helps regulate the serotonin transporter were more likely to become depressed (as measured by either symptoms or disorders) but only when stressed by life events such as employment problems, health issues, housing or relationship issues, and so forth. In addition, childhood maltreatment similarly predicted adult depression only among those with the short allele of the serotonin transporter gene. The implication of this study is that responses to stress and the vulnerability to becoming depressed due to the stress can be predicted by a genetic measure of serotonin function.

Similarly findings can be seen in stressed nonhuman primates. Infant macaque monkeys raised in peer groups were compared to those reared with their mothers. (Lack of normal mothering can be considered an early stress or trauma in both nonhuman primates and in humans.) The biological marker for stress was adrenocorticotropic hormone (ACTH) and cortisol levels, both of which are altered in stressful conditions (see **Adrenocorticotropic Hormone (ACTH), Hypocortisolism and Stress**). Monkeys with the short allele of the serotonin transporter gene showed greater alterations in ACTH and cortisol levels during separation from mothers than those with the longer allele. Thus, these experimental results in monkeys are entirely consistent with the results seen in young adult humans.

Another clue that serotonin may be intimately linked to the effects of stress was demonstrated in a study of infant rhesus monkeys who were raised in mother-reared versus peer-reared conditions. Early maternal deprivation was associated with decreased serotonin metabolites in the cerebrospinal fluid (CSF), a marker for serotonin function in the central nervous system. Furthermore, this decrease in serotonergic function stayed stably lower for the 5 months of the study, implying a potentially long-term effect of stress and trauma on brain serotonin activity.

The regulation of BDNF by specific neurotransmitters is assuredly complex. However, there is some evidence that serotonin plays a role in neurogenesis and may, therefore, be linked with BDNF function. As an example, experimental lesions of the serotonin

system have been shown to decrease neurogenesis in animals. Serotonin agonists increase neurogenesis, an effect blocked by serotonin antagonists. Thus, medications that alter serotonin function, such as SSRIs, can theoretically be expected to increase neurogenesis. Not only do antidepressants enhance BDNF function, but BDNF has been reported to be a potent neurotrophic factor for norepinephrine and serotonin neurons. Thus, if these preliminary data hold true, the relationship between antidepressants and BDNF is best described as reciprocal or bidirectional.

Antidepressants, SSRIs, and Stress Responses

Given that markers of stress and stress-related effects on brain function include both decreasing hippocampal size and diminishing BDNF activity, antidepressants might exert therapeutic effects by preventing or reversing both these effects. Evidence supports both these possibilities.

Antidepressants have been demonstrated to increase neurogenesis in rats. Of note, the antidepressant classes that have shown these positive effects include SSRIs, tricyclic antidepressants, and monoamine oxidase (MAO) inhibitors. The positive effects on neuronal remodeling and neurogenesis seem to require 2–4 weeks of treatment, similar to the time frame of clinical response to these agents in depressed individuals. Electroconvulsive therapy (ECT), with its well-documented antidepressant effects, also increases BDNF expression, as does transcranial magnetic stimulation (TMS), a more recently described antidepressant treatment. The mood stabilizers lithium and valproate, used primarily in the treatment of bipolar disorder, show similar effects. Of note, psychotropic agents from other classes such as antipsychotics and opiates are not associated with increased BDNF activity, implying that the effects may be relatively specific to agents with antidepressant or mood-stabilizing effects.

In addition, as hypothesized, antidepressants also reverse the decreased neurogenesis seen in association with chronic stress and can block this effect of chronic stress when administered preventively. This effect, seen in animal experiments, is entirely consistent with the human evidence that antidepressants administered on a long-term maintenance basis can prevent depressive episodes in individuals vulnerable to them.

It must be acknowledged, however, that even in SSRI studies not all the data demonstrate the salutary effects of these agents on increasing BDNF, neurogenesis, or reversing hippocampal atrophy. In some studies, the expected positive effect is simply not observed. In other rodent studies, the behavioral effects of antidepressants were seen but the enhancement of BDNF function was not, implying that BDNF enhancement was not necessary for the antidepressant behavioral effects.

Interestingly, one factor other than antidepressants that can increase neurogenesis/cellular remodeling in the hippocampus is exercise. In animals, exercise plus antidepressants enhanced BDNF function more than antidepressants. This may partially explain the mild antidepressant effect of aerobic exercise in humans.

Similar observations supporting the role of SSRIs in reversing stress-induced changes in the central nervous system have been made with anxiety models in animals and anxiety disorders in humans. Maternally deprived animals with increased acoustic startle responses and enhanced activity of the hypothalamic-pituitary-adrenal (HPA) axis showed a reversal of these effects with the SSRI paroxetine over a 3- to 4-week time frame (similar to the expected time to clinical effect in anxiety disorders in humans.) In humans, in one study, paroxetine both improved symptoms of PTSD but also enhanced cognitive performance and increased hippocampal volume.

Although all antidepressants may ultimately demonstrate the type of positive effect on reversing and/or preventing stress responses, there may be some differences across antidepressant classes. As one example, although some studies are positive, there are a number of other studies suggesting that tricyclic antidepressants do not consistently increase BDNF, as do other antidepressant classes such as SSRIs.

Consistent with the notion that there may be differences across antidepressant classes in modulating stress responses, a recent study compared the effect of an SSRI to reboxetine, an antidepressant that enhances norepinephrine (but not serotonin) in normal individuals responding to positive and negative images. Although both antidepressants diminished responses to negative emotional material, only the SSRI (citalopram) abolished the increased startle response in the context of negative affective images. The implication of this observation is that antidepressants with strong effects on serotonin, such as SSRIs, may be able to selectively block reactions to negative affective situations that may therefore enhance individuals' capacity to resist the negative effects of stress by blunting the affective and biological response to it.

Because virtually all the research discussed in this chapter is recent, only further studies, many of which are ongoing, will establish how consistently antidepressants reverse stress-related brain changes and whether these effects are common to all antidepressant classes and not just to SSRIs.

See Also the Following Articles

Adrenocorticotropic Hormone (ACTH); Antidepressant Actions on Glucocorticoid Receptors; Cytokines, Stress, and Depression; Depression Models; Glucocorticoids, Effects of Stress on; Glucocorticoids, Role in Stress; Hippocampus, Overview; Depression, Immunological Aspects; Primates: Rearing and Effects of Stress on Primate CNS Function; Hypocortisolism and Stress; Stress and CNS Arousal; Genomic Contributions.

Further Reading

Campbell, S., Marriott, M., Nahmias, C., et al. (2004). Lower hippocampal volume in patients suffering from depression: a meta-analysis. *American Journal of Psychiatry* **161**, 598–607.

Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301**, 386–389.

Duman, R. S. (1998). Novel therapeutic approaches beyond the serotonin receptor. *Biological Psychiatry* **44**, 324–335.

Duman, R. S., Heninger, G. R. and Nestler, E. J. (1997). A molecular and cellular theory of depression. *Archives of General Psychiatry* **54**, 597–606.

Duman, R. S., Nakagawa, S. and Malberg, J. (2001). Regulation of adult neurogenesis by antidepressant treatment. *Neuropsychopharmacology* **25**, 836–844.

Harmer, C. J., Shelley, N. C., Cowen, P. J., et al. (2004). Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition. *American Journal of Psychiatry* **161**, 1256–1263.

Hashimoto, K., Shimizu, E. and Iyo, M. (2004). Critical role of brain-derived neurotrophic factor in mood disorders. *Brain Research Reviews* **45**, 104–114.

Kramer, P. D. (2005). *Against depression*. New York: Viking Penguin.

Nemeroff, C. B., Bremner, J. D. and Foa, E. B. (2006). Posttraumatic stress disorder: a state-of-the-science review. *Journal of Psychiatric Research* **40**, 1–21.

Russo-Neustadt, A. A. and Chen, M. J. (2005). Brain-derived neurotrophic factor and antidepressant activity. *Current Pharmaceutical Design* **11**, 1495–1510.

Self-Esteem, Stress, and Emotion

D Roger
University of Canterbury, Christchurch, New Zealand

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D Roger, volume 3, pp 412–416, © 2000, Elsevier Inc.

The Concept of Self-Esteem
Stress and Personality
Stress and Self-Esteem
Conclusion

Glossary

Emotional inhibition	The inhibition of the expression of experienced emotion.
Factor analysis	A statistical technique for identifying clusters of related items (factors) in questionnaires and for identifying redundant items in preliminary questionnaires. Deciding how many factors to extract is usually based on either an eigenvalue-1 criterion or a scree test, with the latter providing the more conservative estimate.
Hypothalamic-pituitary-adrenal (HPA) axis	The link between the hypothalamus, the pituitary, and the adrenal cortex and medulla. The activation of the adrenal medulla in response to demands, including perceived stress, results in elevated secretion of catecholamines such as epinephrine and norepinephrine. Elevated secretion of adrenal corticosteroids such as cortisol also occurs in these circumstances via the release of corticotropic hormones from the pituitary gland.
Neuroticism	The tendency toward emotional lability or emotional sensitivity.
Personality traits	Enduring, stable predispositions to behave in a particular way.
Rumination	The tendency to persist in thinking about emotionally upsetting events that have happened in the past or might occur in the future; also called emotional rehearsal.
Self-esteem	An evaluative measure of how individuals feel about themselves; includes perceived attractiveness, competence, and confidence.

States	Unstable aspects of personality, such as moods, that vary according to circumstances.
Type A behavior pattern	(TABP) Behavior, such as hostile competitiveness and time urgency, that is thought to increase individuals' susceptibility to coronary heart disease.

Self-concept is a central theme in personality psychology. For the purpose of controlled studies, it is usually operationalized as self-esteem, and a number of self-report psychometric scales have been developed to assess the degree to which individuals value themselves. These self-evaluative perceptions are a reflection of the way individuals feel about their attractiveness and their competencies, and the concept of self-esteem has an intuitive logic in the sense that almost everything people do is evaluated in terms of how successful or valued they feel they are.

Self-esteem clearly has important psychological implications, particularly for stress and personality, but the research has been hampered by conceptual and methodological problems. Many of the early definitions were overinclusive, and the assessment of self-esteem has been marred by psychometric shortcomings. However, work over the past 2 decades has begun to address these issues, leading to a clearer and more precise understanding of self-esteem.

The Concept of Self-Esteem

As early as 1890, William James described self-esteem as being dependent on the match between aspirations (which he called pretensions), on the one hand, and achievements or successes, on the other; the greater the discrepancy between aspiration and actual achievement, the lower self-esteem is likely to be. This suggests a constantly changing process, but self-esteem can also be defined as a relatively permanent predisposition or trait. Indeed, James acknowledged that people tend to have a more or less consistent view of their own self-esteem, developed over time and relatively unaffected by circumstance. Other early theories on the self-concept suggested that the opinions of significant others constituted a kind of mirror in which we view and evaluate ourselves. Although this social self-esteem might vary according to the criticism or the praise we receive, these perceptions became generalized to form an enduring composite of self-attitudes.

Recent trends in research on the self have revived the distinction between relatively stable trait self-esteem and unstable state self-esteem and have provided evidence to show that the distinction is empirically meaningful. For example, assessing self-esteem repeatedly over successive days yields groups with relatively stable or unstable self-esteem. When these two groups are given either positive or negative feedback on their social skills, they differ in their responses, with the unstable high self-esteem subjects generally having more favorable responses to positive feedback but more defensive reactions to negative feedback. The results of these studies suggest that people with unstable high self-esteem are in fact characterized by a poorly developed self-concept and are more reliant on external evaluations and opinions than those with stable high self-esteem.

If the distinction between stable and unstable tendencies were not made, all of these subjects would have had apparently high self-esteem, and the questionnaire responses of the unstable subjects masked an underlying uncertainty and lack of confidence. This approach has been echoed in research on perceived control, which at the extremes of a normal distribution distinguishes between subjects with either an internal or an external locus of control. The former consider their behavior to be a consequence of their own actions, whereas the latter typically attribute events in their lives to luck, chance, or powerful others. However, a further distinction can be drawn between congruent and defensive externals – those who genuinely believe in fate or luck (congruents) and those who say they do in order to protect themselves in the event of failure (defensives). The defensive subjects are identified by their scores on measures of social desirability, which are thought to reflect a desire for approval.

In the context of self-esteem, a similar distinction can be drawn between subjects with high self-esteem and a high need for approval and those with high self-esteem but a low need for approval, based on their scores on social desirability. When these groups are asked to perform ambiguous tasks that allow believable success and failure feedback to be given and are then invited to indicate what they would like their ideal performance to be on a subsequent task, the subjects whose ideal is most diminished by the failure feedback on the analogous task are those with defensive high self-esteem. By contrast, high self-esteem subjects with a low need for approval are unaffected by such experimental manipulations.

More recent research has shown that the degree to which self-esteem is affected by feedback is moderated by other psychological variables, including emotional intelligence (EQ). Self-esteem is positively correlated with measures of EQ, but when individuals with either high or low EQ are exposed to laboratory manipulations that induce negative affect, those with high EQ showed less reduction in self-esteem than those with low EQ.

These effects have been borne out in more naturalistic settings that have incorporated other correlates of self-esteem, such as the degree of perceived support from significant others. A study of distress among cancer patients showed that a perceived lack of social support led to lower self-esteem, which in turn was linked to greater distress. Not surprisingly, self-esteem is also strongly negatively correlated with depression, and a related study of the link between negative marital interactions and depression in arthritis sufferers showed that the relationship was significantly moderated by self-esteem. Conversely, prospective research designs have shown that having social support is associated with high self-esteem, which in turn increases optimism while decreasing depression.

The overall findings confirm that the construct does indeed have two facets: a relatively stable, resilient self-regard, and a more changeable aspect that depends on external circumstances. Although people with genuinely high underlying self-esteem are less affected by a perceived lack of competence in particular situations than those with either low self-esteem or unstable high self-esteem, exposure to self-critical conditions results in diminished self-esteem. An important question raised by these findings is whether self-esteem can systematically be enhanced by targeted training interventions. One study that suggested that it can be involved the random allocation of physically disabled women to two training groups, one of which included a self-esteem enhancement program. The follow-up showed that the women in the experimental group that received the self-esteem training had significantly improved self-esteem as well as lower levels of depression.

However, one of the problems with experiments on self-esteem is the question of reliability. Self-esteem is typically assessed by self-report questionnaires, and it is difficult to ensure that people will respond honestly, particularly where the questions have a personal, evaluative connotation. The problem of response bias is inherent in questionnaire assessment techniques and can only adequately be overcome by using some form of control scale such as social desirability.

Another psychometric problem inherent in scale construction is the distinction between general and multidimensional structure. For example, one of the more widely used self-esteem scales is the 10-item Rosenberg Self-Esteem Inventory. This has the advantage of being extremely brief and easy to administer, but it is usually considered to be unidimensional and it does not break down into different domains of self-esteem. Other popular measures that offer systems for scoring responses into subscales covering areas such as academic, family, and social self-esteem are also thought to represent a single general factor, despite the breakdown into different domains. By contrast, many of the more recent self-esteem scales are predicated specifically on multidimensional models of self-concept. The subscales within these questionnaires typically cover broad domains of physical, social, and academic self-esteem, and there is evidence to show that the subscales are empirically discriminable. However, the most satisfactory of these models are also hierarchical and include a higher-order general factor. Indeed, when conservative techniques such as scree tests are used for extracting factors rather than an eigenvalue-1 criterion, a single general factor is usually indicated for self-esteem questionnaires.

In principle, higher-order general factors in personality are more likely to be independent of one another than lower-order ones. Neuroticism and extraversion, for example, are postulated as statistically independent dimensions. However, self-esteem correlates significantly negatively with neuroticism and positively with extraversion, and it correlates significantly with a wide range of other personality variables as well. This is perhaps not altogether surprising in view of the degree of item-overlap among them. Neuroticism scales are a case in point. They include items that are explicitly concerned with self-esteem, such as "Are you troubled with feelings of inferiority?" and "Do you worry a lot about your looks?"

Setting aside the effects of item-overlap, it may be that the patterning of these other personality dimensions is, in fact, determined by individual differences in self-esteem – in other words, that self-esteem might represent a highest-order construct. What is clear from the research findings is that self-esteem exerts a powerful influence on behavior, and this can be captured by a compensatory model that formalizes the relationship between self-esteem and other aspects of personality, with particular reference to stress and health.

Stress and Personality

Stress has been defined in a variety of ways, with each definition having important implications for the way in which it is measured. Stimulus definitions, for example, led to the development of life-event scales for assessing stress. These instruments continue to be used, despite widespread evidence that life-event reporting is fundamentally flawed. By contrast, response definitions have often focused on physiological changes, and this has proved to be a particularly useful avenue of research for linking stress and illness. Recent advances in psychoneuroimmunology have shown that sustained activation of the HPA axis following exposure to stress may have deleterious consequences; the elevation in adrenal catecholamines

has a significant impact on cardiovascular function, and there is good evidence for the role of adrenal corticosteroids in compromising immune competence.

However, wide individual differences in stress reactions suggest that physiological response mechanisms alone provide only a partial answer. A number of researchers proposed that aspects of personality might play a significant role, and early work on the concept of hardiness suggested that having an internal locus of control may serve as a buffer against stress. Hardiness is a composite of three dimensions: control, commitment, and challenge. Early research had suggested that hardiness, and particularly the control component, might be important in buffering individuals against stress. Unfortunately, subsequent findings using locus of control to investigate the role of personality in stress and illness have been equivocal, as have findings with other variables such as neuroticism and extraversion.

Locus of control is a dimension of personality, but the stimulus properties of control can also be investigated by manipulating the degree of control available over environmental factors. Pioneering experiments in the context of stress used pairs of monkeys, one having control over shock avoidance (the executive monkey) and a yoked companion monkey that did not. The findings suggested that the executive monkey was more likely than the yoked monkey to develop gastrointestinal lesions through anticipatory stress arousal. However, other researchers obtained the opposite result, and it has been pointed out that the relationship between control and stress is more complex; control tends to be harmful primarily when it involves engaging in active and effortful coping responses. The picture becomes even more complicated when individual differences in perceived control are included; having an internal locus of control has been shown to be associated with a greater susceptibility to stress among subjects who experienced high levels of uncontrollable life events.

In view of the inconclusive findings from research on the role of personality in stress and illness, it has been argued that the personality constructs used in earlier studies were inappropriate because they had not been developed specifically in the context of stress research. A number of scales for assessing aspects of emotional response styles have emerged from stress research programs, and subsequent validation studies of these scales have shown that behavior such as rumination and emotional inhibition is strongly related to physiological indices of stress. Rumination in particular has been shown to prolong heart-rate recovery and corticosteroid secretion, and recent findings have shown a significant link between rumination and suppressed immune function. Other work has implicated the tendency to inhibit the expression of emotion in a range of physiological response mechanisms, including delayed muscle tension recovery following exposure to stress. Research on emotion control is thus set within the context of adrenomedullary and adrenocortical response mechanisms mediated by personality, and the findings suggest a new definition of stress as a preoccupation with emotional upset.

Apart from emotional inhibition and rumination, other important factors with implications for stress include coping styles and the type A behavior pattern (TABP). It has been pointed out that there is confusion in the literature over the precise definition of coping, but coping inventories have typically isolated three coping domains involving rational (or problem-focused), emotional, and avoidance strategies. More recent research has uncovered a fourth domain entitled detachment within a four-factor questionnaire, and factor analyses of the four-factor structure has subsequently combined the detached and emotional components into a single bipolar scale that significantly moderates the impact of adaptational stress.

In the same way, a more detailed analysis of TABP has resulted in a reformulation of the construct as a normally distributed dimension of achievement orientation. This work has followed from earlier inconclusive attempts to inadequately distinguish between toxic and nontoxic components of TABP. The more recent research has yielded questionnaires comprising discrete, inversely correlated toxic and nontoxic subscales, in which the former is characterized by hostile competitiveness and the latter by more cooperative achievement styles.

Research with the new scales has shown a consistent pattern of intercorrelations among them and suggested a broader category of toxic behavior that might draw the different strands together. This composite risk index of toxic behavioral tendencies (rumination and emotional inhibition, a tendency towards toxic achievement styles, and a reliance on emotional rather than detached coping) is a potentially powerful tool for predicting deterioration in health status during adaptation, but it does not offer a clear understanding of why some individuals behave in a toxic fashion. Placing the findings on personality and stress in the context of self-esteem offers a way of explaining the findings.

Stress and Self-Esteem

Researchers have speculated that rumination may be the result of factors such as gender or a lack of social support. More fundamentally, however, ruminating about emotional upset may represent an attempt to

preserve self-esteem. TABP has, in fact, been interpreted in this way, by describing type As as having rigid and unrealistic ways of evaluating their own self-worth based on unrealistically high expectations of themselves. A failure to meet these expectations leads to low self-esteem, which these individuals then attempt to compensate for by engaging in the hard-driving, competitive, and aggressive behavior that characterizes TABP. If compensatory behavior is associated with toxic achievement motivation, a positive association might be expected between self-esteem and nontoxic achieving, and this has indeed been found to be the case.

These findings have recently been extended in research examining the costs of attempting to salvage self-esteem when unrealistic expectations are not met. These studies have focused on the goals of efforts to maintain self-esteem and have shown that individuals with purely self-validating goals respond to threats to self-worth in ways that affect both mental and physical health. Clearly, high self-esteem has positive effects on health only when it is pursued in ways that are not contingent on self-validation, a finding that echoes the distinction made earlier between trait and state self-esteem. Noncontingent self-esteem corresponds to a relatively enduring sense of self-worth that is not dependent on others' opinions. Similarly, problem-focused coping strategies are ineffective for highly self-critical perfectionists, but are effective for those with low self-criticism; in the former group, high achievement motivation is undermined by self-criticism, which in turn compromises coping efforts.

Self-esteem has also been shown to act as a buffer against both anxiety and depression. High self-esteem plays a pivotal role in providing protection against anxiety provoked by perceived threats, and in a study that used measures of both optimism and self-esteem to predict postpartum depression, optimism was associated with less depression up to 2 weeks postpartum, but the association with self-esteem was still evident 4 weeks later. Even when the effects of optimism and initial levels of depression were parcelled out, the association between self-esteem and postpartum depression remained.

The stress-buffering effects of self-esteem have been directly tested in experimental studies that used both psychological and physiological measures of stress. These experiments compared groups of subjects who differed in their self-esteem, and the results showed that subjects with high self-esteem suffered less distress, performed better on a cognitive task, and showed less physiological arousal in response to laboratory stressors. The physiological measure in these studies was heart-rate elevation and recovery, but other research has used direct measures of immune competence to investigate the effects of self-esteem. For example, assays of natural killer (NK) cell activity showed that NK cytotoxicity was significantly suppressed as a consequence of negative self-evaluation.

Conclusion

In conclusion, the evidence implicating self-esteem in human stress responses is unequivocal. Low self-esteem is significantly associated with deleterious psychological, physiological, and immunological consequences, whereas high self-esteem serves a protective or buffering function, and these findings hold true even when the effects of other personality factors are parcelled out. Recent research suggests that self-esteem may act as a superordinate construct that determines the patterning of a wide range of lower-order factors. This model suggests that low self-esteem may govern a range of stress-related personality factors, including emotional inhibition and rumination, the tendency to engage in toxic patterns of achievement behavior such as hostile competitiveness, and the attribution of control externally rather than internally. It is argued that these patterns of behavior are used to compensate for the negative feelings resulting from low self-esteem and that engaging in them serves to perpetuate a maladaptive cycle that becomes self-reinforcing.

Further Reading

Bracken, B. A. (ed.) (1996). Handbook of self-concept: developmental, social, and clinical considerations. New York: John Wiley.

Crocker, J. and Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin* **130**, 392–414.

Hughes, R. B., Robinson-Whelan, S., Taylor, H. B., et al. (2004). Enhancing self-esteem in women with physical disabilities. *Rehabilitation Psychology* **49**, 295–302.

Rector, N. and Roger, D. (1997). The stress buffering effects of self-esteem. *Personality and Individual Differences* **23**, 799–808.

Roger, D. (1999). The role of cognitive rumination, coping styles and self-esteem in moderating adaptational responses to stress. In: Mervielde, I., Deary, I., De Fruyt, F. & Ostendorf, F. (eds.) *Personality psychology in Europe* (vol. 7), pp. 351–364. Tilburg: Tilburg University Press.

Roger, D. and Najarian, B. (1998). The relationship between emotional rumination and cortisol secretion under stress. *Personality and Individual Differences* **24**, 531–538.

Schroeder, D. H. and Costa, P. T. (1984). Influence of life event stress on physical illness. *Journal of Personality & Social Behaviour* **46**, 853–863.

Steptoe, A. (1983). Stress, helplessness and control: the implications of laboratory studies. *Journal of Psychosomatic Research* **27**, 361–367.

Thomsen, D. K., Mehlsen, M. Y., Hokland, M., et al. (2004). Negative thoughts and health: associations among rumination, immunity, and health care utilization in a young and elderly sample. *Psychosomatic Medicine* **66**, 363–371.

Self-Harm *See:* Suicide, Biology of; Eating Disorders and Stress.

Selye, Hans

B Tuchweber and P Bois
University of Montreal, Montreal, Quebec, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by G Jasmin and P Bois, volume 3, pp 417–418,
© 2000 Elsevier Inc.

Overview
Education and Awards
Selye's Message

Overview

Hans Selye (**Figure 1**), born in Vienna in 1907, is internationally acknowledged as the father of stress and, arguably, as the greatest physician of all time. Selye's work on stress has had an enormous impact on biomedical sciences and the practice of medicine. His extraordinary energy resulted in the publication during his lifetime of more than 1450 scientific papers and 22 books, of which 15 are monographs that have been translated into many languages. It is noteworthy that in the past 3 years, more than 20 years after his death in 1982, his publications have received more than 695 citations. His concern about gaining public understanding of this work led him to write several books for the layman, some of which gained great popularity, such as *The stress of life* and *Stress without distress*. Selye was an accomplished linguist able to present his work in many parts of the world, where he was eagerly sought as a speaker. As a lecturer and teacher he inspired countless listeners and students worldwide.

Selye was driven by an insatiable inquisitiveness for the pursuit of truth. Well served by a legendary memory, a great mind, and intuitive skills, he understood the importance of reflection following experimentation and gathered in his laboratories the brain power that was necessary to challenge established truths and generate new paradigms.

Education and Awards

Selye's scientific education began in 1924 at the German University of Prague, where he obtained a

Figure 1 Hans Selye.

PhD in chemistry and a doctorate in medicine. As was customary for medical students at that university, he did part of his training abroad, at the University of Paris and the University of Rome. Selye undertook postdoctoral studies in Prague and then was nominated assistant in experimental pathology in charge of the histology laboratory. In 1931, he was awarded a Rockefeller Research Fellowship to work at Johns Hopkins University in Baltimore. A year later he moved to McGill University in Canada, where he was nominated lecturer and assistant professor in the Department of Biochemistry. In 1937 he joined the Department of Histology, where he was nominated associate professor in 1941. It was in 1945 that Selye moved to the Université de Montreal, where he founded and directed the Institut de Médecine et de Chirurgie expérimentales. He held the position of professor and director until his retirement in 1976. Afterwards, Selye founded and became president of the International Institute of Stress in Montreal as well as professor emeritus until his death in 1982.

Selye was the recipient of many honorary degrees and awards from many countries, including 17 honorary degrees from universities in the United States and abroad, 72 medals and awards (e.g., the American Academy Achievement Golden Award, International Kittay, Killam Award), and 12 diplomas. He was elected Fellow of the Royal Society of Canada in 1941 and in 1968 was appointed Companion of the Order of Canada.

Selye's Message

What was Selye's message? Although the starting points were different, the central theme of all his research from 1936 onward was stress and adaptation to stress. His landmark article published in 1936 in *Nature* titled 'A syndrome produced by diverse noxious agents' described the general adaptation syndrome in response to stress with its three stages (alarm, resistance, exhaustion). Very early during his investigations, Selye recognized the significance of the adrenal gland in adaptation and also demonstrated the importance of the pituitary-adrenocortical axis in the stress reaction. He classified the steroids and named the products of the adrenal cortex corticoids, which he subdivided into glucocorticoids and mineralocorticoids. The key role of glucocorticoids in the treatment of inflammatory and immune disorders was also demonstrated.

It is of interest that in Selye's institute, two postdoctoral fellows, Claude Fortier and Roger Guillemin, initiated work on the mechanisms of hormonal regulation of the pituitary-adrenocortical axis. Guillemin later identified the presence of the corticotropin-releasing factor (CRF) in the hypothalamus that stimulated the secretion of adrenocorticotropin (ACTH) by the pituitary gland. He was awarded the Nobel Prize in 1977 for this significant contribution to the understanding of the mechanisms of the stress response.

Further studies by Selye lent support to the view that diseases of adaptation were maladies in which derangements of the adaptation syndrome play a key role. This included the demonstration of hormonal conditioning, the influence of hormones upon reactivity, factors influencing the development of cardiac necroses and related lesions, hormonal production of hypertension, nephrosclerosis, and generalized collagen diseases.

Selye's basic studies were conducted using traditional physiological techniques in animals but were followed by many investigations by some of his graduate students and others with sophisticated techniques including neurochemistry and molecular biology. Selye did not pay much attention to these methods, which he called *in vitro* approaches. Instead, he put his trust in nature, "alone capable," as he used to say, of replying to his questions. He wrote a book on this subject, entitled *In vivo: The case for supramolecular biology,* in which he made a point of his generalist approach.

Selye was one of the great men who were not limited by the bounds of a single discipline, in his case physiology, but who covered several fields (endocrinology, immunology, metabolism) and created new disciplines (neuropsychoimmunology, neuroendocrinology). His work had an impact on human development, cardiovascular disease, central nervous system depression, and the physiology of mind–body interactions. His article on the stress response published in *Nature* in 1936 has been recognized as a psychiatry classic because it led to the study of the effects of stress and hormones, mainly corticosteroids on brain function. Selye acquired worldwide fame through his description of the physiological and biological effects of stress. Although he worked only with animals and never conducted clinical research on or treated humans, the body of knowledge he generated has had a profound impact on our understanding of the consequences of stress on human development, modulation of immune and endocrine response, and psychobiology. There is no doubt that millions of people have benefited and are benefiting today from the concepts he developed.

See Also the Following Article

Alarm Phase and General Adaptation Syndrome.

Further Reading

Selye, H. (1936). A syndrome produced by diverse nocuous agents. *Nature* **138**, 32.
Selye, H. (1947). *Encyclopedia of endocrinology*. Montreal: Acta Inc.
Selye, H. (1956). *The stress of life*. New York: McGraw-Hill.
Selye, H. (1964). *From dream to discovery*. New York: McGraw-Hill.
Selye, H. (1967). *In vivo: The case of supramolecular biology*. New York: Liveright.
Selye, H. (1971). *Hormones and resistance*. New York: Springer-Verlag.
Selye, H. (1976). *Stress in health and diseases*. Boston, MA: Butterworth.
Selye, H. (1979). *The stress of my life, a scientist's memoirs*. New York: Van Nostrand Reinhold.

Separation, Marital *See:* Divorce, Children of; Domestic Violence; Life Events Scale; Life Events and Health; Male Partner Violence; Marital Conflict; Marital Status and Health Problems; Marriage.

Sepsis, Acute Respiratory Distress Syndrome, and Glucocorticoid Resistance

G U Meduri
University of Tennessee Health Science Center
Department of Medicine, Memphis, TN USA

© 2007 Elsevier Inc. All rights reserved.

Introduction
Acute Respiratory Distress Syndrome
Systemic Inflammation-Associated Glucocorticoid Resistance
Nuclear Factor κB and Glucocorticoid Receptor α Interaction
Systemic Inflammation-Induced Glucocorticoid Resistance Can Be Improved with Prolonged Glucocorticoid Supplementation
Findings of Randomized Trials
Conclusion

Glossary

Hypoxemic respiratory failure	A life-threatening reduction in blood oxygenation requiring assisted mechanical ventilation.
Multiple organ dysfunction	Systemic inflammation-associated impairments of vital organs: brain, lung, cardiovascular system, renal, liver, and hematopoietic system.
Sepsis	Systemic inflammation induced by infection.
Systemic inflammation	The release of inflammatory cytokines in the systemic circulation manifesting clinically with fever and increased heart rate, respiratory rate, and white cell blood count.

Introduction

Hundred of thousands of American are affected every year by severe sepsis (700,000) and acute respiratory distress syndrome (ARDS; 190,000) creating a significant short-term and long-term burden on public health and health-care economics. Dysregulated systemic inflammation with persistent elevation of circulating inflammatory cytokines over time is the pathogenetic mechanism for the dysfunction and failure of the vital organs, the leading cause of death in patients with sepsis and ARDS. In sepsis and ARDS, the downregulation of inflammation is *conditio sine qua non* to their resolution. An improved understanding of the cellular mechanisms that initiate, propagate, and restrict inflammation in sepsis and ARDS is essential to make advances on present treatment modalities and to develop more reliable methods of monitoring disease progression and responses to therapy.

It is now appreciated that at cellular level transcription factors nuclear factor (NF-)-κB, which is

activated by inflammatory signals, and the glucocorticoid receptor (GR)α, which is activated by endogenous or exogenous glucocorticoids (GCs), have diametrically opposed functions (stimulatory vs. inhibitory) in regulating inflammation. NF-κB is recognized as the principal driver of the inflammatory response, being responsible for the transcription of more than 100 genes, including tumor necrosis factor (TNF-)α and interleukin(IL-)1β (Figure 1). Once activated, NF-κB and GRα can mutually repress one another through a protein–protein interaction that prevents their DNA binding and subsequent transcriptional activity. Known interactions between NF-κB and the activated glucocorticoid receptor (GC-GRα) are depicted in Figure 1. The activation of one transcription factor in excess of the binding (inhibitory) capacity of the other shifts cellular responses toward the increased (dysregulated) or decreased (regulated) transcription of inflammatory mediators over time (Figure 2). The role of GC

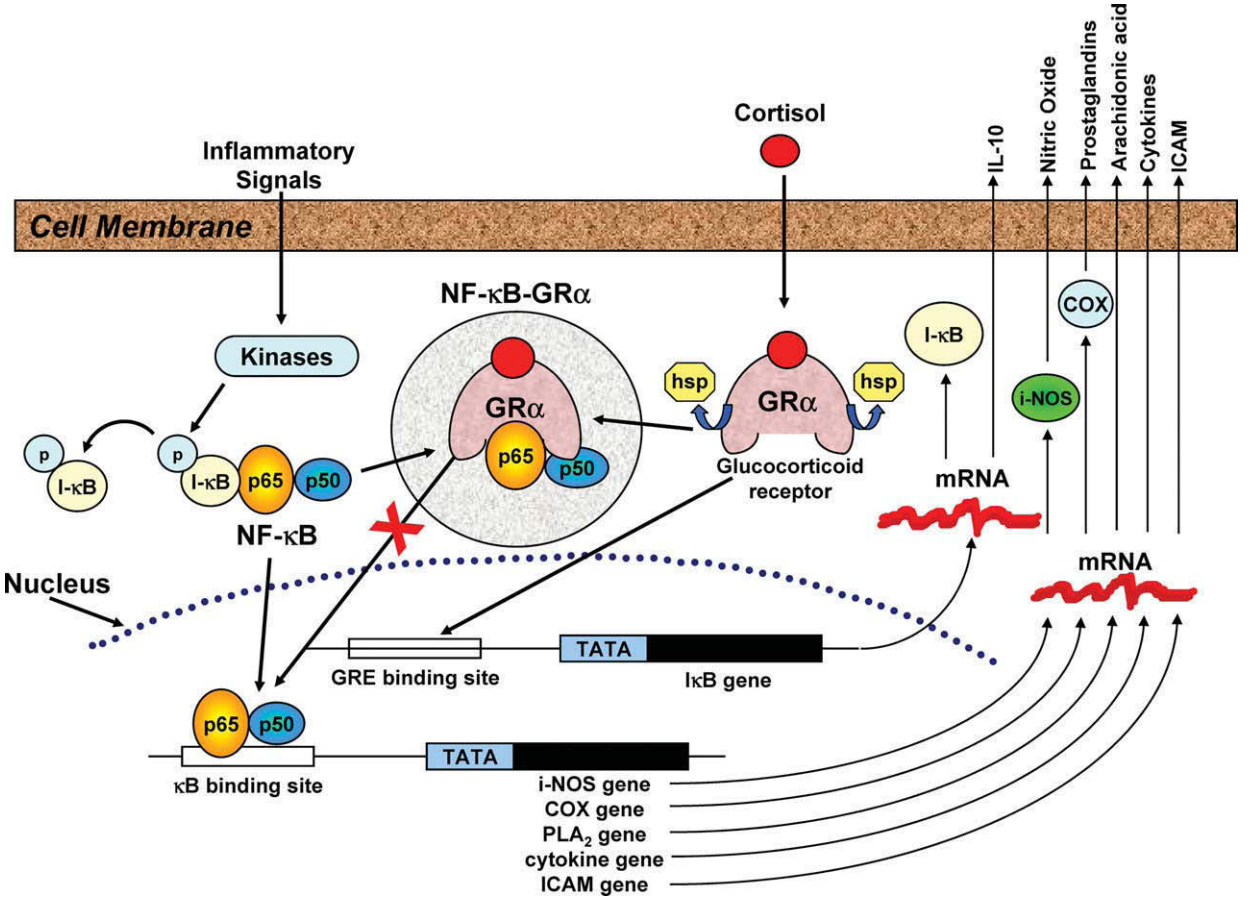


Figure 1 Positive regulatory loop interaction between NF-κB and the activated glucocorticoid receptor (GC-GRα). When cells are stimulated by inflammatory signals, specific kinases phosphorylate the inhibitory protein IκB and cause its rapid degradation. The activated form of NF-κB then moves to the nucleus, initiating the transcription of mRNA of inflammatory cytokines, chemokines, cell adhesion molecules, and inflammation-associated enzymes (cyclooxygenase, phospholipase A2, and inducible nitric oxide). Cortisol or exogenous glucocorticoids freely cross into the cytoplasm and bind to their specific glucocorticoid receptors (GRα) to form GC-GRα. GC-GRα complexes may influence NF-κB activity in five major ways: (1) physically interacting with the p65 subunit to form an inactive (GC-GRα–NF-κB) complex, (2) inducing the synthesis of the inhibitory protein IκBα via interaction with glucocorticoid-responsive element DNA in the promoter of the IκB gene, (3) blocking the degradation of IκBα via enhanced synthesis of IL-10, (4) impairing the TNF-α-induced degradation of IκBα, and (5) competing for limited amounts of GR co-activators such as cAMP response element binding protein (CREB)-binding protein (CBP) and steroid receptor coactivator (SRC-)1. IκBα, in addition to holding NF-κB in an inactive cytoplasmic state, also translocates into the nucleus, where it binds activated NF-κB complexes to induce their export to the cytoplasm. GC-GRα may also decrease the stability of mRNA of several inflammatory cytokines and other molecules. The products of the genes that are stimulated by NF-κB activate this transcription factor. Thus, TNF-α and IL-1β both activate and are activated by NF-κB by forming a positive regulatory loop that amplifies and perpetuates inflammation. Reproduced with permission from Meduri, G. U., Muthiah, M. P., Carratu, P., et al. (2005), Nuclear factor-kappaB- and glucocorticoid receptor alpha-mediated mechanisms in the regulation of systemic and pulmonary inflammation during sepsis and acute respiratory distress syndrome: evidence for inflammation-induced target tissue resistance to glucocorticoids, *Neuroimmunomodulation* **12**(6), 321–338.

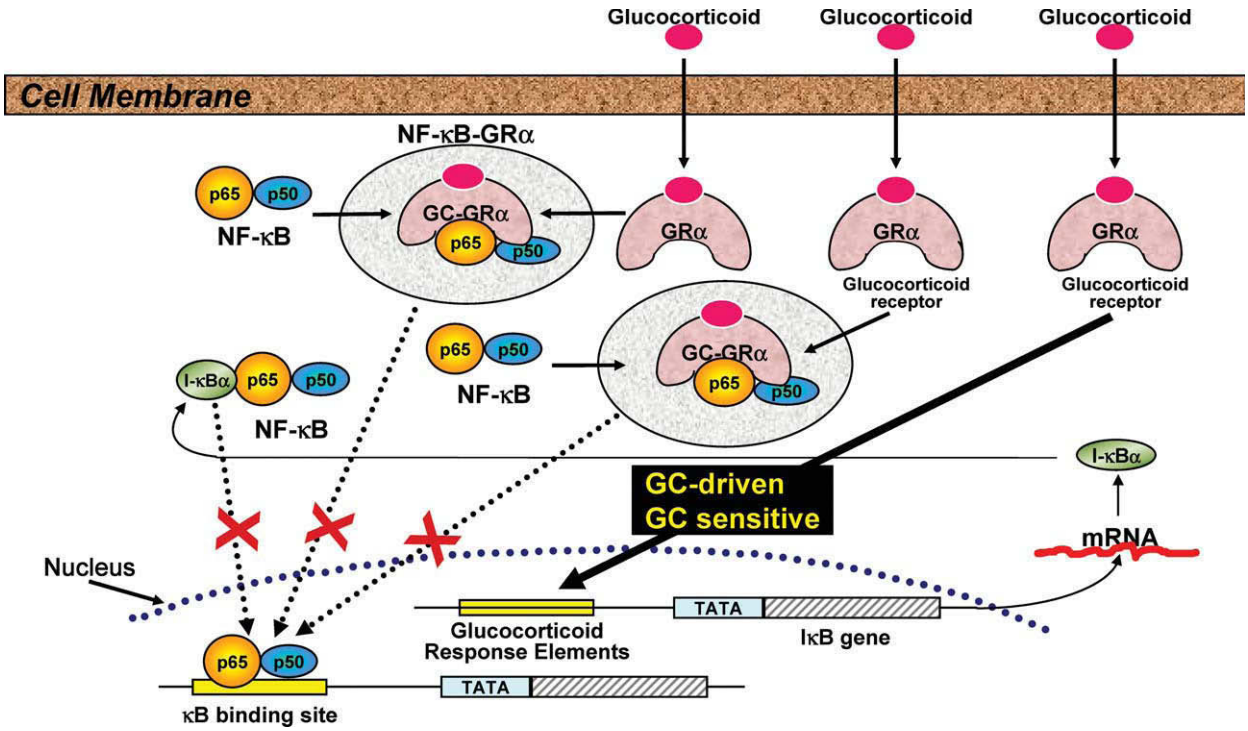


Figure 2 Regulated response interaction between NF-κB and the activated glucocorticoid receptor (GC-GRα). GC-GRα activation sufficient to maintain NF-κB levels in homeostasis and achieve a reduction in the transcription of inflammatory mediators over time is shown. In improvers, both GC-GRα binding to NF-κB and nuclear GC-GRα binding increased significantly over time, indicating an excess activation of GC-GRα compared to NF-κB (GC-GRα-driven response). Reproduced with permission from Meduri, G. U., Muthiah, M. P., Carratu, P., et al. (2005), Nuclear factor-kappaB- and glucocorticoid receptor alpha-mediated mechanisms in the regulation of systemic and pulmonary inflammation during sepsis and acute respiratory distress syndrome: evidence for inflammation-induced target tissue resistance to glucocorticoids, *Neuroimmunomodulation* **12**(6), 321–338.

supplementation in patients with life-threatening systemic inflammation is undergoing a cyclical reassessment, stirred by a new understanding of the role of GCs in regulating inflammation and the positive results of recent randomized controlled studies.

Acute Respiratory Distress Syndrome

ARDS is a life-threatening form of hypoxemic respiratory failure (mortality of 30–60%) that develops within 12–48 h of a patient's acquiring an acute illness complicated by severe systemic inflammation. The most common condition precipitating ARDS is severe sepsis, and the mortality rate for sepsis-induced ARDS has not significantly decreased in the last 20 years. In ARDS, the evolution (regulated vs. dysregulated) of systemic inflammation in the first week of mechanical ventilation determines the physiological progression of the disease and the final outcome. Patients failing to improve in the first week of mechanical ventilation (nonimprovers) have, compared to improvers, a persistent elevation in circulating and bronchoalveolar lavage (BAL) levels of inflammatory cytokines and chemokines, markers of alveolocapillary membrane (ACM) permeability and fibrogenesis. In sepsis and ARDS, the downregulation of inflammation is essential for the restoration of homeostasis and survival.

Systemic Inflammation-Associated Glucocorticoid Resistance

GCs as end effectors of the hypothalamic-pituitary-adrenal (HPA) axis are the most important physiological inhibitors of inflammation, affecting hundreds of genes involved in stress-related homeostasis. However, endogenous GCs are not always effective in suppressing life-threatening systemic inflammation, and the degree of cortisolemia frequently correlates with the severity of illness and the mortality rate. Unquestionably, the elevation of GC secretion in nonsurvivors of sepsis and ARDS is inadequate to meet the needs of the concurrent inflammatory response and its adverse systemic effects. Its failure to suppress inflammation could be due to tissue resistance to GCs, the inadequacy of the level and duration of endogenous GC elevation to suppress a systemic inflammatory response gone awry, or both.

The concept of acquired GC resistance was first introduced by Kass and Finland in 1957. These investigators suggested that increased blood cortisol levels in nonsurvivors of sepsis might reflect a blocking of steroidal activity or transport as a consequence of the infection. In this situation, a small increase in blood levels with a low (equal to or less than physiological) dose of exogenous GCs was believed to be sufficient to facilitate the passage of steroids into host cells. GC resistance was originally described as a primary inherited familial syndrome and was recently recognized as an acquired condition. Among others, acquired immune tissue-specific GR resistance has been described in patients with asthma, acquired immunodeficiency syndrome (AIDS), severe sepsis, and sepsis-induced ARDS.

In vitro studies showed that cytokines may induce resistance to GCs by reducing GR binding affinity to cortisol and/or GC response elements (GREs). Such abnormalities of GR function were demonstrated in T cells incubated with a combination of interleukin (IL-)2 and IL-4, IL-1, IL-6, and interferon (IFN)-γ, or IL-13. GC resistance was induced in a cytokine concentration-dependent fashion and was reversed by the removal of cytokines. GR-mediated resistance in the presence of systemic inflammation was also studied in experimental models of sepsis and sepsis-induced ARDS. In a sheep model of sepsis-induced ARDS, the maximal binding capacity of GR decreased continuously after endotoxin infusion, whereas there was a marked elevation of cortisol levels. The reduced GR binding correlated negatively ($\mathbf{r} = -0.87$, $p < 0.01$) with phospholipase A_2 (PLA_2) activity, a gene that is stimulated by NF-κB. In a rat model of septic shock, GR blockade by mifepristone (RU 486) exacerbated the physiological and pathological changes induced by endotoxemia. PLA_2 activity in rats with 80% GR blockade was more marked than in those with 50% GR blockade. The monocytes of patients with sepsis developed near-total GC resistance in vitro when cytokines, especially IL-2, were added.

Nuclear Factor κB and Glucocorticoid Receptor α Interaction

New evidence indicates that a primary pathogenetic mechanism explaining GC resistance or insensitivity in patients with life-threatening systemic inflammation is the imbalance between excessive NF-κB activation by inflammatory signals and deficient GRα activation by endogenous or exogenous glucocorticoids. Using an *ex vivo* model of systemic inflammation, our group investigated intracellular upstream and downstream events associated with DNA binding of NF-κB and GRα in naïve peripheral blood leukocytes (PBLs) stimulated with longitudinal plasma specimens obtained from 28 patients with sepsis-induced ARDS and correlated the physiological progression with laboratory findings. Exposure of naïve cells to longitudinal plasma samples led to divergent directions in NF-κB and GRα activation that reflected the severity of systemic inflammation (defined by plasma TNF-α and IL-1β levels). The activation of one transcription factor in excess of the other shifted cellular responses toward decreased (GRα-driven) or increased (NF-κB-driven) transcription of inflammatory mediators over time (Figures 2–3).

Plasma samples from patients with declining inflammatory cytokine levels over time elicited a progressive increase in all measured aspects of GC-GRα-mediated activity ($p = 0.0001$) and a corresponding reduction in NF-κB nuclear binding ($p = 0.0001$) and transcription of TNF-α and IL-1β (regulated, GRα-driven response; Figure 2). In contrast, plasma samples from patients with sustained elevations in inflammatory cytokine levels over time elicited only modest longitudinal increases in GC-GRα-mediated activity ($p = 0.04$) and a progressive increase in NF-κB nuclear binding over time ($p = 0.0001$) that was most striking in nonsurvivors (dysregulated, NF-κB-driven response; Figure 3). These findings demonstrate that insufficient GC-GRα-mediated activity is an important mechanism for early loss of autoregulation (i.e., downregulation of NF-κB activation) in nonimprovers. Deficient GRα activity in naïve cells exposed to plasma from patients with dysregulated inflammation was observed despite elevated circulating cortisol and adrenocorticotropic hormone (ACTH) levels, implicating inflammatory cytokine-driven excess NF-κB activation as an important mechanism for target-organ insensitivity and/or resistance to cortisol.

Our findings are in agreement with two longitudinal studies that investigated NF-κB binding activity directly in the peripheral blood mononuclear cells of patients with sepsis or trauma. In both studies, nonsurvivors (compared to survivors) had a progressive increase in NF-κB activity over time. In one longitudinal study, nonsurvivors of septic shock had by days 2–6 a 200% increase in NF-κB activity compared to day 1. Similarly, using our *ex vivo* model we found that NF-κB binding activity on day 3 of ARDS clearly separated patients by outcome. In an immunohistochemistry analysis of lung tissue, areas with severe (vs. mild) ARDS had a higher nuclear uptake of NF-κB ($13 + 1.3$ vs. $7 + 2.9$; $p = 0.01$) and a lower ratio of GRα : NF-κB nuclear uptake (0.5 ± 0.2 vs. 1.5 ± 0.2; $p = 0.007$). Thus, an uninhibited increase in NF-κB activation in circulating cells over time

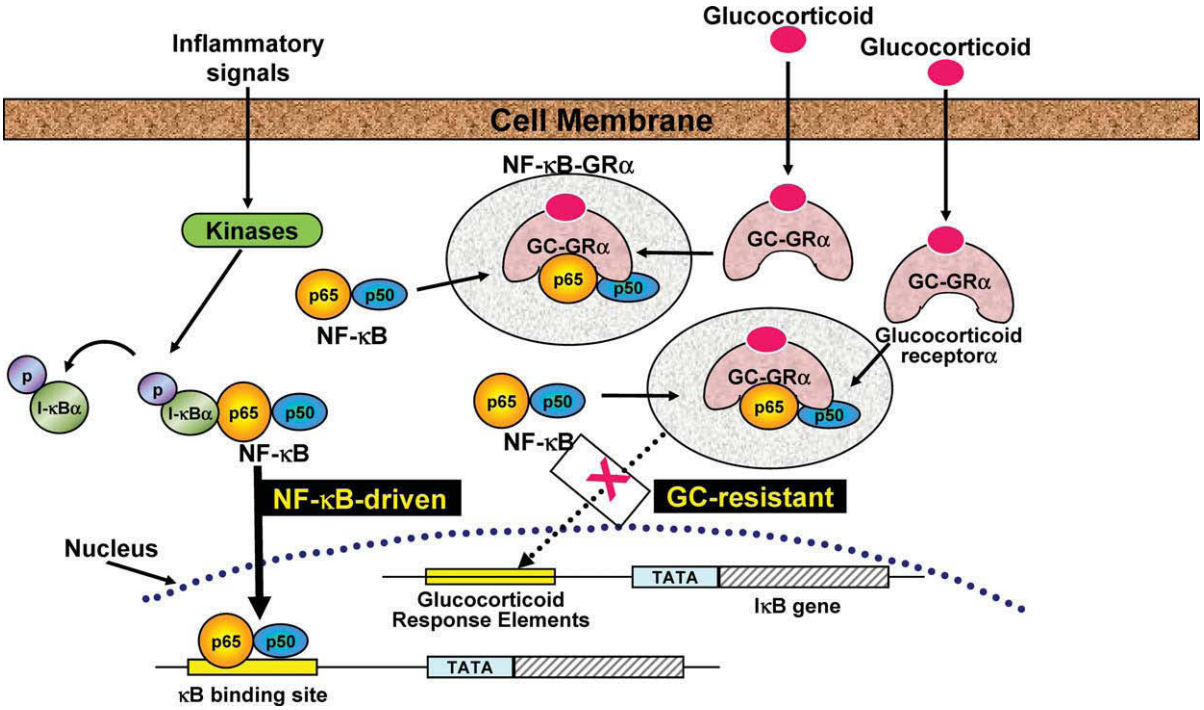


Figure 3 Dysregulated response interaction between NF-κB and the activated glucocorticoid receptor (GC-GRα). An excess of NF-κB activation is shown, leading to the protracted transcription of inflammatory mediators over time. In nonimprovers, GC-GRα binding to NF-κB was modestly increased, whereas nuclear NF-κB binding increased substantially over time and nuclear GC-GRα and cytoplasmic IκBα levels declined (NF-κB-driven response). Reproduced with permission from Meduri, G. U., Muthiah, M. P., Carratu, P., et al. (2005), Nuclear factor-kappaB- and glucocorticoid receptor alpha-mediated mechanisms in the regulation of systemic and pulmonary inflammation during sepsis and acute respiratory distress syndrome: evidence for inflammation-induced target tissue resistance to glucocorticoids, *Neuroimmunomodulation* **12**(6), 321–338.

appears to be a significant premortem pathogenetic component of lethal sepsis and ARDS.

Systemic Inflammation-Induced Glucocorticoid Resistance Can Be Improved with Prolonged Glucocorticoid Supplementation

Our findings place GC-GRα-mediated downregulation of NF-κB activity as a critical factor for the reestablishment of homeostasis during the acute, life-threatening systemic inflammation that accompanies sepsis and ARDS. A subgroup of 11 patients with dysregulated inflammation and deficient GC-GRα activity were randomized to prolonged GC supplementation initiated on ARDS day 9 ± 3. Quantitatively adequate and prolonged hormonal supplementation restored GR anti-inflammatory function, decreasing the production of inflammatory cytokines, inflammatory cytokine-driven HPA axis activity, and inflammatory cytokine-driven organ dysfunction. Normal PBL exposed to plasma samples collected during GC versus placebo treatment exhibited rapid, progressive, and significant increases in GRα-mediated activities (GRα binding to NF-κB, GRα binding to GRE DNA, stimulation of inhibitory protein IκBα, and stimulation of IL-10 transcription) and significant reductions in NF-κB κb binding and transcription of TNF-α and IL-1β. During GC treatment, the relationship between NF-κB and GC-GRα signaling pathways changed from an initial NF-κB-driven and GC-GRα-resistant state to a GC-GRα-driven and GC-GRα-sensitive one. In aggregate, the findings indicate that systemic inflammation-induced GC resistance is an acquired generalized process mediated by excess NF-κB activation and potentially reversible by increasing GC-GRα activation with quantitatively adequate and prolonged GC supplementation.

Findings of Randomized Trials

At present, the administration of low-dose GCs is the only treatment shown in randomized trials to have a significant positive biological and physiological effect across the longitudinal spectrum of the severity of sepsis-associated systemic inflammation: severe sepsis, septic shock, early ARDS, and unresolving ARDS.

Table 1 Prolonged glucocorticoid treatment in sepsis and ARDS: effect on mortality[a]

Severity of sepsis	*Number of RCTs*	*Number of patients (death of total)*			*Source*[b]
		Treated	*Control*	*p value*	
Severe sepsis	1	0 of 23 (0%)	7 of 23 (30%)	0.009	5
Septic shock	5	7 of 22	12 of 19		2
		3 of 20	4 of 20		3
		6 of 23	10 of 21		4
		82 of 151	91 of 149		1
		8 of 20	12 of 20		10
Total from meta-analysis		106 of 236 (45%)	129 of 229 (56%)	0.01	8
Early ARDS	1	50 of 63 (21%)	16 of 28 (43%)	0.03	6
Unresolving ARDS	2	2 of 16 (12%)	5 of 8 (62%)	0.03	7
		32 of 90 (36%)	31 of 90 (34%)	NS	9

[a]ARDS, acute respiratory distress; RCTs, randomized controlled trials.
[b]Data from:
1. Annane, D., Sebille, V., Charpentier, C., et al. (2002), Effect of treatment with low doses of hydrocortisone and fludrocortisone on mortality in patients with septic shock, *Journal of the American Medical Association* **288**(7), 62–871.
2. Bollaert, P. E., Charpentier, C., Levy, B., et al. (1998), Reversal of late septic shock with supraphysiologic doses of hydrocortisone, *Critical Care Medicine* **26**(4), 645–650.
3. Briegel, J., Forst, H., Haller, M., et al. (1999), Stress doses of hydrocortisone reverse hyperdynamic septic shock: a prospective, randomized, double-blind, single-center study, *Critical Care Medicine* **27**(4), 723–732.
4. Chawla, K., Kupfer, Y., Goldman, I., et al. (1999), Hydrocortisone reverses refractory septic shock, *Critical Care Medicine* **27**, A33.
5. Confalonieri, M., Urbino, R., Potena, A., et al. (2005), Hydrocortisone infusion for severe community-acquired pneumonia: a preliminary randomized study, *American Journal of Respiratory Critical Care Medicine* **171**(3), 242–248.
6. Meduri, G. U., Golden, E. , Freire, A. X., et al. (2005), Methylprednisolone infusion significantly improves lung function in patients with early acute respiratory distress syndrome: results of a randomized controlled trial, *Chest* **128**, S129.
7. Meduri, G. U., Headley, S., Golden, E., et al. (1998), Effect of prolonged methylprednisolone therapy in unresolving acute respiratory distress syndrome: a randomized controlled trial, *Journal of the American Medical Association* **280**, 159–165.
8. Minneci, P. C., Deans, K. J., Banks, S. M., et al. (2004), Meta-analysis: the effect of steroids on survival and shock during sepsis depends on the dose, *Annals of Internal Medicine* **141**(1), 47–56.
9. Steinberg, K. P., Hudson, L. D., Goodman, R. B., et al. (2006), Efficacy and safety of corticosteroids for persistent acute respiratory distress syndrome, *New England Journal of Medicine* **354**(16), 1671–1684.
10. Yildiz, O., Doganay, M., Aygen, B., et al. (2002), Physiological-dose steroid therapy in sepsis, *Critical Care* **6**(3), 251–259.

All randomized studies showed significant physiological improvement and a reduction in the duration of stays in the intensive care unit. The overall effects on mortality are shown in Table 1.

In prolonged low-dose GC treatment, maximal benefits are obtained by (1) avoiding the premature removal of an effective drug, leading to rebound inflammation and physiological deterioration, and (2) preventing potential complications associated with GC treatment (infection surveillance and avoidance of paralysis). Uncontrolled studies have also shown a likely benefit in patients with viral pneumonia-induced ARDS, including SARS. The mortality rate in 849 cases of SARS treated with antiviral therapy and GC supplementation was only 7.7%. In hantavirus pulmonary syndrome, the mortality rate decreased from 47% in patients treated with antiviral therapy to 13% when GC supplementation was also included. The findings for severe viral pneumonia and the results of recent randomized controlled trials (RCTs) for ARDS indicate that prolonged methylprednisolone use might useful in the prevention of morbidity and mortality associated with avian influenza.

Conclusion

Dysregulated systemic inflammation has enormous health-care consequences. A new understanding of the cellular mechanisms responsible for inflammation-associated GC resistance provides a new rationale for investigating prolonged low-dose GC supplementation. The consistent positive results of small- to moderate-size randomized studies are very encouraging. Federal financial support is urgently needed to conduct additional RCTs investigating this inexpensive off-patent treatment in patients with life-threatening systemic inflammation.

Acknowledgments

This work was supported by Grants from the Assisi Foundation of Memphis.

Further Reading

Annane, D., Sebille, V., Charpentier, C., et al. (2002). Effect of treatment with low doses of hydrocortisone and fludrocortisone on mortality in patients with septic

shock. *Journal of the American Medical Association* **288**(7), 62–871.

Annane, D., Sebille, V., Troche, G., et al. (2002). A 3-level prognostic classification in septic shock based on cortisol levels and cortisol response to corticotropin. *Journal of the American Medical Association* **283**(8), 1038–1045.

Bollaert, P. E., Charpentier, C., Levy, B., et al. (1998). Reversal of late septic shock with supraphysiologic doses of hydrocortisone. *Critical Care Medicine* **26**(4), 645–650.

Briegel, J., Forst, H., Haller, M., et al. (1999). Stress doses of hydrocortisone reverse hyperdynamic septic shock: a prospective, randomized, double-blind, single-center study. *Critical Care Medicine* **27**(4), 723–732.

Chawla, K., Kupfer, Y., Goldman, I., et al. (1999). Hydrocortisone reverses refractory septic shock. *Critical Care Medicine* **27**, A33.

Cheng, V. C., Tang, B. S., Wu, A. K., et al. (2004). Medical treatment of viral pneumonia including SARS in immunocompetent adult. *Journal of Infection* **49**(4), 262–273.

Confalonieri, M., Urbino, R., Potena, A., et al. (2005). Hydrocortisone infusion for severe community-acquired pneumonia: a preliminary randomized study. *American Journal of Respiratory Critical Care Medicine* **171**(3), 242–248.

Kass, E. H. and Finland, M. (1957). Adrenocortical hormones and the management of infection. *Annual Review of Medicine* **8**, 1–18.

Meduri, G. U., Golden, E., Freire, A. X., et al. (2005). Methylprednisolone infusion significantly improves lung function in patients with early acute respiratory distress syndrome: results of a randomized controlled trial. *Chest* **128**, S129.

Meduri, G. U., Headley, S., Golden, E., et al. (1998). Effect of prolonged methylprednisolone therapy in unresolving acute respiratory distress syndrome: a randomized controlled trial. *Journal of the American Medical Association* **280**, 159–165.

Meduri, G. U., Muthiah, M. P., Carratu, P., et al. (2005). Nuclear factor-kappaB- and glucocorticoid receptor alpha-mediated mechanisms in the regulation of systemic and pulmonary inflammation during sepsis and acute respiratory distress syndrome: evidence for inflammation-induced target tissue resistance to glucocorticoids. *Neuroimmunomodulation* **12**(6), 321–338.

Meduri, G. U., Tolley, E. A., Chrousos, G. P., et al. (2002). Prolonged methylprednisolone treatment suppresses systemic inflammation in patients with unresolving acute respiratory distress syndrome: evidence for inadequate endogenous glucocorticoid secretion and inflammation-induced immune cell resistance to glucocorticoids. *American Journal of Respiratory Critical Care Medicine* **165**(7), 983–991.

Meduri, G. U. and Yates, C. R. (2004). Systemic inflammation-associated glucocorticoid resistance and outcome of ARDS. *Annals of the New York Academy of Sciences* **1024**, 24–53.

Minneci, P. C., Deans, K. J., Banks, S. M., et al. (2004). Meta-analysis: the effect of steroids on survival and shock during sepsis depends on the dose. *Annals of Internal Medicine* **141**(1), 47–56.

Rhen, T. and Cidlowski, J. A. (2005). Antiinflammatory action of glucocorticoids – new mechanisms for old drugs. *New England Journal of Medicine* **353**(16), 1711–1723.

Steinberg, K. P., Hudson, L. D., Goodman, R. B., et al. (2006). Efficacy and safety of corticosteroids for persistent acute respiratory distress syndrome. *New England Journal of Medicine* **354**(16), 1671–1684.

Yildiz, O., Doganay, M., Aygen, B., et al. (2002). Physiological-dose steroid therapy in sepsis. *Critical Care* **6**(3), 251–259.

Septic Shock *See:* Hypotension, Hypovolemia, and Septic Shock; Sepsis, Acute Respiratory Distress Syndrome and Glucocorticoid Resistance.

Serotonin

B C Williams and Y-C Lo
University of Edinburgh and Western General Hospital, Edinburgh, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 419–423, © 2000, Elsevier Inc.

Serotonin and Adrenocortical Function
Serotonin and Renal Function
Conclusion

Glossary

5-Hydroxytryptamine (5-HT) receptors — Plasma membrane proteins that mediate the biological effects of serotonin through the activation of an intracellular second-messenger system. There are at least seven different classes of 5-HT receptors, which are defined according to their pharmacological properties using selective agonists and antagonists and/or their molecular structure as determined by molecular cloning techniques; some of these classes of 5-HT receptors have been divided further into subclasses. Not all of the 5-HT receptors identified have yet been assigned a clear biological function.

Aromatic-L-amino acid decarboxylase (L-AAAD) — A chemical (EC 4.1.1.28) that catalyzes the decarboxylation of l-5-hyroxytryptophan and l-dopa (l-3,4-dihydroxyphenylalanine) to serotonin and dopamine (3,4-dihydroxyphenylethylamine), respectively. It requires pyridoxal phosphate as a coenzyme and occurs most abundantly in the kidney, liver, gastrointestinal tract, adrenal gland, and brain; previously known as hydroxytryptophan decarboxylase or dopa decarboxylase.

Serotonin (5-hydroxytryptamine, 5-HT) — A biogenic amine (first isolated from serum by Irvine Page in 1948) shown to affect vascular tone. Its other name (adopted by Erspamer in the 1940s) was enteramine because it was found to be associated with enterochromaffin cells of the gastric mucosa. By the 1950s, the chemical structure was established and serotonin/enteramine was renamed 5-hydroxytryptamine.

Zona fasciculata — A broad layer of cells in contact with the zona glomerulosa and forming the main mass of the adrenal cortex. These cells secrete the steroid hormones cortisol (in humans) and corticosterone (in rodents). These steroids act on glucocorticoid receptors, which are distributed very widely in mammalian tissues, and produce many different biological actions, including the breakdown of carbohydrates, proteins, and fats.

Zona glomerulosa — A narrow layer of cells in the outer region of the adrenal gland, situated directly beneath the adrenal capsule, whose specific function is to secrete the steroid hormone aldosterone, which acts on mineralocorticoid receptors in the distal convoluted tubule of the kidney to cause sodium retention and potassium loss.

Serotonin and Adrenocortical Function

In Vitro Effects of Serotonin in the Adrenal Cortex

The stimulatory effect of 5-HT on adrenocortical steroidogenesis in several species, including the rabbit, guinea pig, rat, and human, was discovered by Rosenkrantz in 1959. Based on the sodium-retaining effects and chromatographic properties, the steroid stimulated by 5-HT was identified as aldosterone, which is synthesized in the zona glomerulosa cells of the adrenal cortex. Concurrent experiments demonstrated that rat serum (but not rat plasma) stimulated steroid production in the rat zona glomerulosa, indicating that a component of the blood-clotting system (later identified as 5-HT) was involved. These landmark experiments led to subsequent studies by numerous workers to determine the role of 5-HT in adrenocortical function. Most of these studies were carried out in rat models for the reason that the zona glomerulosa in this species can be cleanly dissected from the zona fasciculata and zona reticularis and, following enzymatic disaggregation of the tissue, the zona glomerulosa cells can be obtained with less than 5% contamination by other cells types. When rats are placed on a low sodium chloride diet, the sensitivity of the aldosterone response to 5-HT is increased in the zona glomerulosa. Conversely, when either a high sodium chloride diet or a low potassium chloride diet is given, the zona glomerulosa cells become less sensitive to 5-HT. This clearly indicates a dependency of the aldosterone response to 5-HT on the sodium/potassium status, which is also observed for other zona glomerulosa-specific stimuli, such as angiotensin II and plasma potassium. In contrast, no

direct steroidogenic response to 5-HT is found in cells of the zona fasciculata or zona reticularis. The stimulation of aldosterone by 5-HT is associated with an increase in cAMP in rat zona glomerulosa cells, and these responses are inhibited by 5-HT antagonists such as methysergide, cyproheptadine, and ketanserin, indicating the presence of 5-HT cell surface receptors in the zona glomerulosa. Although this receptor was identified tentatively as a 5-HT_2 receptor (now renamed the 5-HT_{2A} receptor) based on the effects of ketanserin, subsequent cloning of this receptor did not support this notion. No mRNA for this receptor (which is coupled to phospholipase C) has been found in zona glomerulosa cells, and 5-HT does not stimulate phospholipase C in these cells. The 5-HT receptor in rat zona glomerulosa cells still awaits clear identification. In human zona glomerulosa cells, the 5-HT receptor has been partially characterized, pharmacologically, as a 5-HT_4 receptor (which couples to adenylate cyclase), and the 5-HT_4-selective agonist zacopride is a potent stimulus to aldosterone in human zona glomerulosa tissue. In both rat and human zona glomerulosa cells, 5-HT-stimulated aldosterone secretion is also coupled to Ca^{2+} influx and is inhibited by verapamil, suggesting the involvement of voltage-dependent Ca^{2+} channels.

Ex Vivo Effects of Serotonin in the Adrenal Cortex

In the isolated perfused rat adrenal gland, 5-HT increases the perfusion flow rate, which correlates with the increased secretion of both aldosterone and corticosterone. No 5-HT receptors have been characterized in the adrenal vasculature to date, but because the blood flow rate is an important determinant for steroid secretion *in vivo,* these potential vascular effects of 5-HT in the adrenal cortex could be important physiologically.

In Vivo Effects of Serotonin on Adrenocortical Steroidogenesis

Infusions of high doses of 5-HT in humans cause an increase in aldosterone but not cortisol secretion. Because plasma levels of 5-HT are normally extremely low, due to the active 5-HT reuptake mechanism in the platelets, the potential physiological relevance of this observation is unclear. The administration of the 5-HT precursors L-tryptophan and 5-hydroxytryptophan (5-HTP) in humans produces an increase in both aldosterone and cortisol secretion, and it is thought that these effects are mediated by the activation of the hypothalamic–pituitary–adrenal (HPA) axis because it has been demonstrated clearly in rat tissues *in vitro* that 5-HT can stimulate both corticotropin releasing hormone (CRH) release and adrenocorticotropic hormone (ACTH) release. Intraperitoneal injections of 5-HTP in rats lead to an increase in both aldosterone and corticosterone secretion and plasma renin activity *in vivo*. The corticosterone response to 5-HT can be completely inhibited by the co-administration of dexamethasone, whereas the aldosterone response to 5-HTP is only partially inhibited. The co-administration of the angiotensin-converting enzyme inhibitor captopril partially inhibits the aldosterone response to 5-HTP treatment and does not affect the corticosterone response. These studies therefore indicate that corticosterone (a zona fasciculata steroid) can be regulated by 5-HT *in vivo* by the stimulation of the HPA axis, whereas the aldosterone response is more complex, involving the activation of the HPA axis, the renin–angiotensin system, and, in addition, direct effects in the zona glomerulosa. The effects of 5-HT on renin release are thought to be centrally mediated because 5-HT does not stimulate renin release in the kidney. The notion that the direct effects of 5-HT on zona glomerulosa cells could be an important physiological control mechanism for aldosterone secretion is supported by studies that demonstrate that infusions of the 5-HT_4-selective agonist zacopride in humans increase plasma aldosterone but not plasma cortisol.

Sources of Serotonin in the Adrenal Gland

The physiological source of 5-HT that regulates aldosterone secretion in the zona glomerulosa has not been clearly identified to date. Although 5-HT has been located in the adrenal medulla of some species, including the rat and pig by immunohistochemical techniques, the centripetal blood flow in the adrenal gland from the outer cortex to the inner medulla argues against any significant influence of medullary secretory products on adrenocortical function. In the rat, mast cells (which store 5-HT) have been found in the subcapsular portions of the adrenal artery, and drugs that cause mast-cell degranulation can stimulate aldosterone secretion in the perfused rat adrenal gland. Physiological mechanisms leading to the specific release of 5-HT from these mast cells within the adrenal cortex, however, have not been identified. Other studies have focused on platelets as a possible physiological source of adrenocortical 5-HT. The use of ^{14}C-radiolabeled 5-HT to track the platelet deposition of 5-HT *in vivo* in rats has demonstrated that 5-HT can potentially be released in some endocrine tissues, including the thyroid gland and the adrenal gland, but the release mechanisms are unclear. Perhaps the most likely source of adrenocortical 5-HT is the enzyme L-AAAD, which has been found, by immunohistochemistry, to be specifically associated with the zona glomerulosa in the rat adrenal cortex.

This enzyme converts 5-HTP (which circulates in plasma and is not stored in platelets) to 5-HT and can be inhibited by carbidopa. The incubation of rat adrenal zona glomerulosa tissue *in vitro* with 5-HTP, at physiological concentrations, leads to an increase in aldosterone and 5-HT, and both are completely inhibited in the presence of carbidopa. Furthermore, this enzyme in the zona glomerulosa appears to be regulated by changes in the dietary sodium chloride intake in parallel with observed changes in aldosterone secretion. Thus, the enzyme activity is increased by a low sodium chloride intake and is decreased by a high sodium chloride intake.

Serotonin and Renal Function

Effects of Serotonin on Renal Vasculature

The effects of 5-HT on the renal circulation are variable, depending on the species studied. Most investigations have been carried out in rats or dogs, with few comprehensive studies in humans. The administration of 5-HT to rats *in vivo* (by infusion or intraperitoneal injection) increases renal vascular resistance and reduces the renal blood flow rate. These effects are also observed in the *ex vivo*-isolated perfused rat kidney, are independent of intact renal nerves, and are inhibited by ketanserin, a 5-HT_{2A} antagonist. In the dog, similar effects are observed, although they are clearly more dependent on renal innervation, in that the stimulation of the sympathetic nerves results in a vasodilator response to 5-HT rather than a vasoconstriction. It has been suggested that 5-HT_{1D} receptors may be more important in mediating 5-HT-induced renal vasoconstriction than 5-HT_{2A} receptors in the dog because the 5-HT agonist sumatriptan can reduce renal blood in the presence of a ketanserin blockade. The 5-HT receptor that mediates renal vasodilatation in the dog has not been characterized, although it is thought to be associated with the vascular endothelium and to be coupled to nitric oxide production. Most studies in humans have indicated either no change or only a modest reduction in renal plasma flow and glomerular filtration rate (GFR) in response to intravenous 5-HT. The physiological role of 5-HT in modulating renal hemodynamics in humans is therefore still unclear.

Effects of Serotonin on Renal Tubule Function

In rats, intravenous or intraperitoneal administration of 5-HT causes a decrease in the urine flow rate and a reduction in sodium excretion. Both the antidiuretic and the antinatriuretic effects of 5-HT are thought to involve a direct effect on tubular function, although the antinatriuretic could be partially mediated by 5-HT-stimulated aldosterone secretion. Similarly, in dogs, 5-HT can produce both antidiuresis and antinatriuresis, and these effects are inhibited by the nonselective 5-HT_1 antagonist methysergide. In humans, most studies show that 5-HT infusions lead predominantly to an antinatriuretic effect, with no significant or only modest antidiuresis. Similar responses are observed in response to 5-HTP.

Renal Sources of Serotonin

Although blood platelets store 5-HT, which can be released in certain disease states characterized by intravascular coagulation, it is unlikely that platelet 5-HT contributes to the effects of 5-HT on renal function in normal physiological conditions. Renal tissue, however, can synthesize 5-HT from the precursor 5-HTP via the action of L-AAAD, which is found throughout the renal cortex and also in the renal medulla. Immunostaining for L-AAAD is particularly intense in the renal tubules of the rat kidney, the site of action for 5-HT in relation to its antinatriuretic effect.

Actions of Serotonin Renal Prodrugs

Renal prodrugs are essentially precursors of biologically active endogenous compounds that are preferentially metabolized to the active compound in the kidney. An intravenous infusion of either 5-HTP or the renal prodrug γ-L-glutamyl-5-hydroxy-L-tryptophan (γ-glutamyl-5-HTP) in humans reduces sodium excretion and increases urinary 5-HT excretion due to their metabolism to 5-HT within the kidney. In the presence of carbidopa, which inhibits the activity of L-AAAD, the sodium retention caused by these drugs and their metabolism to 5-HT are completely abolished.

Conclusion

5-HT is a stimulus for both mineralocorticoid (aldosterone) and glucocorticoid (cortisol and corticosterone) secretion. Within the adrenal cortex, 5-HT is a specific stimulus for aldosterone secretion in the zona glomerulosa. In the kidney, the renal hemodynamic effects of 5-HT still require clarification, especially in humans, whereas the antinatriuretic effects of 5-HT are reproducible in several species, including humans. The enzyme L-AAAD is likely to play a key role in the local production of 5-HT in the adrenal cortex and in the kidney. It seems logical that the actions of 5-HT within the adrenal cortex and the kidney are coordinated in order to maintain sodium homeostasis, and it will therefore be important to understand the mechanisms that regulate the activity of L-AAAD in these tissues.

See Also the Following Articles

Adrenal Cortex; Aldosterone and Mineralocorticoid Receptors; Glucocorticoids, Overview.

Further Reading

Adler, S. (1977). Serotonin and the kidney. In: Essman, W. B. (ed.) *Serotonin in health and disease* (vol. 4), pp. 99–137. New York: Spectrum.

Burns, N., Brett, L., Olverman, H. J., et al. (1996). The role of L-aromatic amino acid decarboxylase in serotonin-stimulated aldosterone secretion in response to salt intake. *Endocrine Research* **22**, 577–578.

Collis, M. G. and Vanhoutte, P. M. (1977). Vascular reactivity of isolated perfused kidneys from male and female spontaneously hypertensive rats. *Circulation Research* **41**, 759–767.

Contesse, V., Hamel, C., Lefebvre, H., et al. (1996). Activation of 5-hydroxytryptamine(4) receptors causes calcium influx in adrenocortical cells: involvement of calcium in 5-hydroxytryptamine-induced steroid secretion. *Molecular Pharmacology* **49**, 481–493.

Davies, E., Edwards, C. R. W. and Williams, B. C. (1991). Serotonin stimulates calcium influx in isolated rat adrenal zona glomerulosa cells. *Biochemical and Biophysical Research Communications* **179**, 979–984.

Davies, E., Rossiter, S., Edwards, C. R. W., et al. (1991). Serotoninergic stimulation of aldosterone secretion in the rat *in vivo:* role of the renin-angiotensin system. *Journal of Endocrinology* **130**, 347–355.

Davies, E., Rossiter, S., Edwards, C. R. W., et al. (1992). Serotoninergic stimulation of aldosterone secretion *in vivo:* role of the hypothalamo-pituitary adrenal axis. *Journal of Steroid Biochemistry and Molecular Biology* **42**, 29–36.

Erspamer, V. (1966). Peripheral physiological and pharmacological actions of indolealkylamines. *Handbook of Experimental Pharmacology* **19**, 245–359.

Hinson, J. P., Vinson, G. P., Pudney, J., et al. (1989). Adrenal mast cells modulate vascular and secretory responses in the intact adrenal gland of the rat. *Journal of Endocrinology* **121**, 252–260.

Holmes, M. C., Di Renzo, G., Beckford, U., et al. (1982). The role of serotonin in the control of secretion of corticotropin releasing factor. *Journal of Endocrinology* **93**, 151–160.

Lefebvre, H., Contesse, V., Delarue, C., et al. (1993). Effect of the serotonin-4 receptor agonist zacopride on aldosterone secretion from human adrenal cortex: *in vivo* and *in vitro* studies. *Journal of Clinical Endocrinology and Metabolism* **77**, 1662–1666.

Li Kam Wa, T. C., Freestone, S., Samson, R. R., et al. (1993). A comparison of the renal and neuroendocrine effects of two 5-hydroxytryptamine renal prodrugs in normal man. *Clinical Science* **85**, 607–614.

Li Kam Wa, T. C., Freestone, S., Samson, R. R., et al. (1994). The antinatriuretic action of γ-L-glutamyl-5-hydroxy-L-tryptophan is dependent on its decarboxylation to 5-hydroxytryptamine in normal man. *British Journal of Clinical Pharmacology* **38**, 265–269.

McCubbin, J. W., Kaneko, Y. and Page, I. H. (1962). Inhibition of neurogenic vasoconstriction by serotonin. *Circulation Research* **11**, 74–83.

Muller, J. (1987). *Regulation of aldosterone biosynthesis: physiological and clinical aspects*. New York: Springer-Verlag.

Osim, E. E. and Wyllie, J. H. (1983). Loss of 5-hydroxytryptamine from mammalian circulating platelets. *Journal of Physiology* **340**, 77–90.

Shoji, T., Tamaki, T., Fukui, K., et al. (1989). Renal haemodynamic responses to 5-hydroxytryptamine(5-HT): involvement of the 5-HT receptor subtypes in the canine kidney. *European Journal of Pharmacology* **171**, 219–228.

Stier, C. T., McKendall, G. and Itskovitz, H. D. (1984). Serotonin formation in non-blood perfused rat kidneys. *Journal of Pharmacology and Experimental Therapeutics* **228**, 53–56.

Vinson, G. P., Pudney, J. A. and Whitehouse, B. J. (1985). The mammalian adrenal circulation and the relationship between adrenal blood flow and steroidogenesis. *Journal of Endocrinology* **105**, 285–294.

Whiting, M. V. and Cambridge, D. (1995). Canine renovascular responses to sumatriptan and 5-carboxyamidotryptamine: modulation through endothelial $5HT_1$-like receptors by endogenous nitric oxide. *British Journal of Pharmacology* **114**, 969–974.

Serotonin in Stress

S Kusljic and M van den Buuse
Mental Health Research Institute, Parkville, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

Stress and Psychopathology: Possible Role of Serotonin
Stress, Serotonergic Drugs, and Human Psychopathology
Effect of Stress on Serotonin Parameters
Effect of Serotonergic Drugs on Stress Responses
Conclusion

Glossary

Hypothalamic-pituitary-adrenocortical (HPA) axis	The pathway that regulates glucocorticoid hormones secretion in the periphery.
Raphe nuclei	Groups of brain cells that are the main source of serotonin in the forebrain.
Serotonin (5-hydroxytryptamine, 5-HT)	A widely distributed monoamine that has been involved in almost every function of the brain.

Stress and Psychopathology: Possible Role of Serotonin

Serotonin (5-hydroxytryptamine, 5-HT) is a widely distributed monoamine that has been involved in almost every function of the brain. It is known that serotonin acts through a complex receptor system (5-HT_1–5-HT_7), with the number of receptors showing regional variation in terms of their distribution. The brain serotonergic system originates predominantly from cell groups located in either the dorsal or the median raphe nuclei, but some cell bodies are located in the ventrolateral region of the midbrain. Serotonergic neurons have a major influence on the regulation of neuroendocrine function, cardiovascular and respiratory activity, sleep, pain sensitivity, and mood. Serotonin furthermore acts as a growth factor during embryogenesis, and serotonin receptor activity forms a crucial loop in the pathway leading to changes in brain structure. There is extensive evidence for a role of serotonin in human illnesses associated with disrupted brain functioning.

Similarly, stress is involved in many aspects of human disease. At least part of this involvement could be caused by or associated with effects of stress on the serotonergic system and its widespread functions. For example, some of the physiological changes associated with the stress response include enhanced cardiovascular output and respiration, modulation of the immune response, increased cerebral perfusion, and mobilization of energy to maintain brain function. Also more subtle changes in behavioral functioning caused by stress, e.g., arousal, anxiety, and cognition, may involve brain serotonin. Psychological stress, furthermore, can have an important role in the onset and course of mental illness, including depression, anxiety, and schizophrenia.

The influence of stressful life events in humans has received widespread attention. Exposure to early life stress is associated with neurobiological changes in children and adults, which may underlie the increased risk of psychopathology. There is substantial evidence that early life events influence serotonergic function in the brain as well as brain development and subsequent adult behavior. For example, the effects of early traumatic events on the developing brain and the biological abnormalities found in subjects with schizophrenia include hyperactivity of the hypothalamic-pituitary-adrenocortical (HPA) axis and serotonin abnormalities, and structural changes to the brain such as damage to the hippocampus, a brain region enriched in serotonin receptors. Stress accelerates both the release and turnover of brain serotonin. This can lead to restoration of homeostasis, whereas chronic stress can induce a long-lasting imbalance in central nervous neurotransmitter systems, including serotonin. Abuse and neglect during early childhood is linked to low levels of serotonin and risk for hostility, aggression, substance abuse, and suicide.

Several other lines of evidence have linked brain serotonin activity with human psychopathology. A polymorphism in the promoter region of the serotonin transporter gene has been suggested to contribute to anxiety in humans, and the presence of an allele encoding a short sequence in this promoter region increases the risk for the development of a depressive disorder. Studies in a population of rhesus monkeys have shown that the short sequence allele corresponds to low concentrations of the serotonin metabolite 5-hydroxyindolacetic acid in the cerebrospinal fluid. Mutations in the gene for the enzyme tryptophan hydroxylase may also predispose to mood disturbances. Alterations in the several synaptic proteins, transcription factors, and proteins involved in neuronal growth/differentiation that are modified in their expression in experimental models of chronic stress can be reversed by treatment with antidepressants, thus suggesting an important role for the serotonin system.

Stress, Serotonergic Drugs, and Human Psychopathology

Evidence for the importance of the serotonin system in depression and anxiety disorders has increased substantially over the recent years, and stress plays a major role in the development of these disorders. The first effective antidepressants (tricyclic antidepressants and monoamine oxidase inhibitors) such as imipramine and iproniazid relied at least in part on their ability to increase serotonin and norepinephrine levels in the synapse. Research on the biological effects of these drugs led to the hypothesis of a deficiency in both these cerebral monoamines in depression. Antidepressants act not only on depressive symptoms but also on anxiety. Most of the evidence suggests that depletion of serotonin in the brain increases anxiety in humans. Thus, drugs that increase the levels of serotonin have anxiolytic effects. A new generation of antidepressants that selectively inhibit the neuronal reuptake of serotonin (selective serotonin reuptake inhibitors [SSRIs]) specifically focused on serotonin neurotransmission with little or no activity on noradrenergic mechanisms. Several drugs, including fluoxetine, citalopram, and paroxetine, had the same clinical efficacy, yet appeared to be better tolerated and lack most of the adverse side effects of tricyclic antidepressants and monoamine oxidase (MAO) inhibitors. Originally introduced for the treatment of depression, the SSRIs have been used as the first-line treatment for a broad range of psychiatric disorders, including anxiety, obsessive-compulsive disorder, social phobia, and posttraumatic stress disorder. The target of SSRIs is the presynaptic serotonin transporter protein, which is responsible for clearance of synaptic and plasma serotonin.

Effect of Stress on Serotonin Parameters

Stress increases concentrations of serotonin and its metabolites in several brain regions indicating increased turnover rates of the neurotransmitter. Moreover, stress increases the activity of tryptophan hydroxylase, the biosynthetic enzyme confined to serotonin-producing cells. Therefore, studies of the effect of stress on serotonergic function in laboratory animals may lead to a better understanding of the biological bases of psychopathology.

Repeated exposure of rats to forced swimming increased serotonin concentrations in the striatum but reduced serotonin levels in the lateral septum. Chronic restraint stress in rats revealed increased serotonin turnover in the hippocampus, whereas chronic social defeat in mice upregulated expression of the MAO-A gene in the raphe nuclei. MAO-A is an isoenzyme that preferentially oxidizes serotonin. An upregulated expression of the MAO-A gene accelerates serotonin degradation and is associated with enhanced expression of the serotonin transporter.

Furthermore, early life stress such as maternal deprivation can affect several aspects of serotonin transmission later in life. For example, maternal deprivation at postnatal day 3 has long-term effects on the serotonergic responsiveness of CA1 pyramidal neurons in the hippocampus of the rat. Adult CA1 hippocampal cells from maternally deprived rats compared to nondeprived rats show a significantly attenuated response to serotonin. This attenuation is also seen in adult rats subjected to high doses of exogenously administered corticosterone or to chronic stress.

Corticosterone regulates hippocampal physiology, particularly in the CA1 region, and the 5-HT_{1A} receptor effector pathway. However, prolonged elevation of the corticosteroid concentration does not alter 5-HT_{1A} receptor mRNA levels in pyramidal neurons. In addition, in rats that were subjected to unpredictable stressors for 3 weeks and that displayed increased adrenal weight and attenuated body weight compared with controls, the 5-HT_{1A} receptor mRNA expression was not changed after chronic stress exposure. Therefore, stress may alter serotonergic function possibly by posttranslational modification of the specific serotonin receptor subtype or changes in the serotonin receptor signal transduction pathway.

Effect of Serotonergic Drugs on Stress Responses

Serotonin and HPA Axis Activity

The primary neuroendocrine response to acute stress includes increased secretion of cortisol in humans or corticosterone in rodents. Glucocorticoid hormones are the final effectors of the HPA axis that participate in the control of homeostasis and the response of the body to stressors. Corticosterone acts on both the high-affinity mineralocorticoid receptors and the lower affinity glucocorticoid receptors. Mineralocorticoid receptors are found in some limbic brain structures that are also rich in serotonin innervation, such as the hippocampus, whereas glucocorticoid receptors are found in several brain regions, including frontal cortex and hypothalamic paraventricular nucleus. Glucocorticoids play a key regulatory role in the termination of the stress response by exerting negative feedback at the levels of hypothalamus and pituitary, and there is increasing evidence to link serotonin with HPA regulation. Serotonin-containing neurons directly innervate corticotropin-releasing hormone (CRH)-containing cells located in the paraventricular

nucleus of the hypothalamus. Serotonergic innervation of the paraventricular nucleus modulates the release of CRH, leading to the release of adrenocorticotropic hormone (ACTH), which triggers glucocorticoid secretion from the adrenal cortex. Furthermore, administration of the serotonin precursors L-tryptophan and 5-hydroxytryptophan stimulates ACTH and cortisol release, as does the serotonin releasing agent fenfluramine. 5-HT_{1A} and 5-HT_{2A} receptors are the main serotonin receptors mediating serotonergic stimulation of the HPA axis. Activation of 5-HT_{1A} and 5-HT_{2A} receptors increases plasma corticosterone levels. Conversely, the stimulatory effect of serotonin on CRH secretion was antagonized by the 5-HT_1 and 5-HT_2 receptor antagonist metergoline and by the selective 5-HT_2 receptor antagonists ketanserin and ritanserin. Furthermore, administration to rats of the 5-HT_{1A} agonist, 8-hydroxy-2-(di-n-propylamino) tetralin (8-OH-DPAT), resulted in activation of the HPA axis as evidenced by increased plasma ACTH and corticosterone concentrations. Chronic administration of 8-OH-DPAT resulted in increased CRH concentrations in the piriform cortex, entorhinal cortex, amygdala, and hippocampus. However, administration of 8-OH-DPAT had no effect on the 5-HIAA/5-HT ratio. In contrast, 8-OH-DPAT treatment was able to decrease plasma corticosterone responses following acoustic stress, conditioned fear, and cocaine administration. This indicates that the predominant effect of 8-OH-DPAT on adrenocortical responses is mediated at postsynaptic sites not involved in the regulation of cerebral biogenic amine metabolism.

Acute administration of SSRIs stimulates the HPA axis, whereas continuous treatment desensitizes it. The immediate effect of SSRIs on the HPA axis function occurs in the hypothalamic paraventricular nucleus via postsynaptic serotonin receptors on CRH neurons. It is likely that serotonin 5-HT_{1A} and 5-HT_{2A} postsynaptic receptors are indirectly stimulated by SSRI administration.

Physiological Responses: Blood Pressure/Body Temperature

A typical neuroendocrine response also involves the increased secretion of catecholamines (epinephrine and norepinephrine) from the sympathetic nervous system and adrenal medulla. Central catecholamine neurons play a critical role in cardiovascular responses to stress-related stimuli. Hypertension and increased heart rate are key components of the cardiovascular response to stress, leading to the essential increase in tissue perfusion that has to occur in stressful situations. Among the complex interactive systems that contribute to the central regulation of cardiovascular function, many areas are under the direct influence of serotonergic pathways. Both blood pressure and heart rate may be increased or decreased by the activation of distinct brain serotonin receptors. For example, pharmacological activation of central 5-HT_3 receptors evokes a significant hypotensive effect in rats. By contrast, activation of 5-HT_2 receptors in the lateral hypothalamus and in the paraventricular nucleus increases blood pressure and heart rate. Intraventricular injection of the 5-HT_{2C} agonist (1-(3-chlorophenyl)piperazine) (mCPP) elicited a significant increase in blood pressure in nonstressed animals. However, injection of mCPP did not change the hypertensive or heart rate responses induced by restraint stress. Furthermore, cardiovascular effects caused by the central administration of mCPP were blocked by pretreatment with the 5-HT_{2C} antagonist SDZ SER 082. Conversely, administration of SDZ SER 082 blunted stress-induced hypertension without modifying stress-induced increase in heart rate. Several serotonergic drugs with full or partial 5-HT_{1A} receptor agonist activity, such as 8-OH-DPAT, buspirone, and MDL 73,005EF, were shown to reduce both blood pressure and heart rate responses to open-field novelty stress (**Figure 1**). This effect was seen without changes in the exploratory locomotor activity response in this test.

Administration of serotonergic drugs may affect other physiological responses associated with stress, such as body temperature changes. Acute administration of 3,4-methylenedioxymethamphetamine (MDMA, ecstasy) is associated with rapid hyperthermia that can last several hours and is linked to monoamine release from nerve endings in the brain. In addition to hyperthermia, other, often fatal, physiological problems occur, including tachycardia, renal failure, edema, and thrombocytopenia. MDMA acts by releasing monoamines, including serotonin, but the receptor mechanisms underlying MDMA-induced hyperthermia and cutaneous vasoconstriction remain unclear. MDMA-induced hyperthermia is associated with reduced cutaneous blood flow, which may contribute to the hyperthermia by diminishing normal heat transfer from the body to the environment. The specific 5-$HT_{2A/2C}$ agonist DOI also causes marked hyperthermia, suggesting the possibility that stimulation of 5-$HT_{2A/2C}$ receptors might cause sympathetically mediated cutaneous vasoconstriction. The 5-HT_{2A} receptor antagonist ketanserin reversed the cutaneous vasoconstriction induced by DOI in rabbits and tail vasoconstriction and hyperthermia in rats. Furthermore, administration of the antipsychotic clozapine reverses hyperthermia and sympathetically mediated cutaneous vasoconstriction induced by MDMA in rabbits and rats. These mechanisms

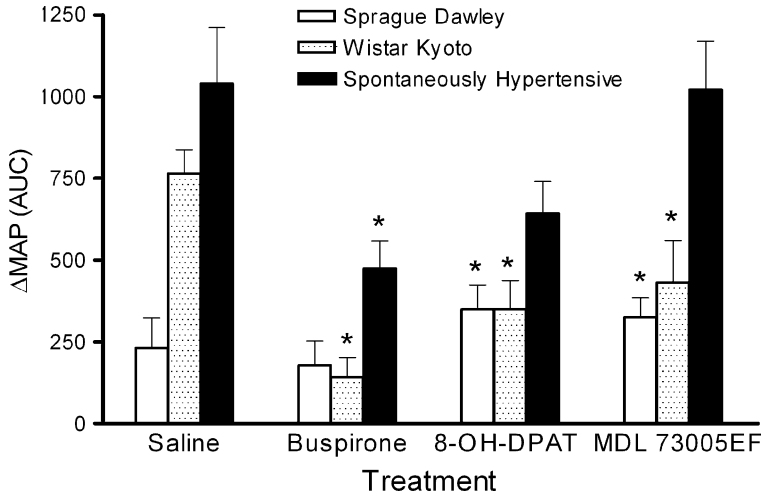


Figure 1 Blood pressure increases caused by open-field novelty stress are modulated by serotonergic drugs in Sprague-Dawley rats, Wistar-Kyoto rats (WKY) and spontaneously hypertensive rats (SHR). Pretreatment with 0.5 mg/kg of the 5-HT_{1A} receptor partial agonist and dopamine D_2 receptor antagonist, buspirone, reduced the blood pressure increases in WKY and SHR. Pretreatment with 0.25 mg/kg of the 5-HT_{1A} receptor full agonist, 8-OH-DPAT, enhanced the blood pressure increase in Sprague Dawley rats, but reduced it in WKY, with no significant change in SHR. This same pattern of strain differences in responding was seen with pretreatment with the 5-HT_{1A} receptor partial agonist, MDL 73,005EF. The rats were placed in 90 cm open field while blood pressure was measured with radiotelemetry. For more data and details, see Van den Buuse and Wegener (2005).

support a role of serotonin mechanisms in stress-induced changes in body temperature. In addition, clozapine and ketanserin, by interactions with 5-HT receptors or by other serotonergic mechanisms, may be of therapeutic benefit in humans suffering from life-threatening hyperthermia such as that induced by ingestion of MDMA.

Conclusion

Serotonin is a widely distributed monoamine that has been involved in almost every function of the brain. Stress influences 5-HT activity, and this interaction could be involved in the development of stress-related illnesses, including depression and schizophrenia. Conversely, serotonergic drugs modulate activity of the HPA axis and physiological responses to stress. Thus, there is a direct reciprocal relationship between stress and activity of the serotonin system in the brain.

See Also the Following Articles

Adrenocorticotropic Hormone (ACTH); Blood Pressure; Corticosteroid Receptors; Depression and Manic-Depressive Illness; Disease, Stress Induced; Glucocorticoids, Effects of Stress on; Glucocorticoids, Overview; Glucocorticoids, Role in Stress; Hypothalamic-Pituitary-Adrenal; Psychological Stressors, Overview; Schizophrenia; Serotonin; Selective Serotonin Reuptake Inhibitors (SSRIs).

Further Reading

Buwalda, B., Kole, M. H., Veenema, A. H., et al. (2005). Long-term effects of social stress on brain and behavior: a focus on hippocampal functioning. *Neuroscience and Biobehavioral Reviews* **29**, 83–97.

Carrasco, G. A. and Van de Kar, L. D. (2003). Neuroendocrine pharmacology of stress. *European Journal of Pharmacology* **463**, 235–272.

Dinan, T. G. (1996). Serotonin and the regulation of hypothalamic-pituitary-adrenal axis function. *Life Sciences* **58**, 1683–1694.

Flügge, G., Van Kampen, M. and Mijnster, M. J. (2004). Perturbations in brain monoamine systems during stress. *Cell and Tissue Research* **315**, 1–14.

McEwen, B. S. (2003). Early life influences on life-long patterns of behavior and health. *Mental Retardation and Developmental Disabilities Research Reviews* **9**, 149–154.

Saphier, D. and Welch, J. E. (1995). Effects of the serotonin$_{1A}$ agonist, 8-hydroxy-2-(di-n-propylamino)tetralin on neurochemical responses to stress. *Journal of Neurochemistry* **64**, 767–776.

Van den Buuse, M. and Wegener, N. (2005). Involvement of serotonin$_{1A}$ receptors in cardiovascular responses to stress: a radio-telemetry study in four rat strains. *European Journal of Pharmacology* **507**, 187–198.

Serotonin Transporter Genetic Modifications

K Sugden
Social, Genetic and Developmental Psychiatry Centre, London, UK

© 2007 Elsevier Inc. All rights reserved.

The Serotonin Transporter
The *5-HTT* Gene
Functional Polymorphisms of the *5-HTT* Gene
The Serotonin Transporter-Linked Polymorphic Region and Stress

Glossary

Enhancer	A short region of DNA that can be bound with proteins (such as transcription factors) to enhance transcription levels of genes. An enhancer does not necessarily have to be located close to the genes it affects, although some enhancers are found within intronic sequences.
Messenger RNA (mRNA)	The primary sequence transcribed from the DNA template. mRNA has intronic sequences (i.e., sequences not represented in the transcript exiting the nucleus) removed by splicing and encodes the amino acid sequence of a protein.
Polyadenylation	The cleavage of mRNA and addition of 50–200 adenosine residues, a poly-A tail. Polyadenylation is an important post-transcriptional modification that ensures the appropriate termination of the molecule at the 3′ end and helps in both the export of the mRNA from the nucleus and protection from exonucleases.
Polyadenylation site	The point of polyadenylation of newly transcribed mRNA.
Serotonin (5-hydroxytryptamine, 5-HT)	A protein present in many tissues, especially blood and nervous tissue; it stimulates a variety of smooth muscles and nerves and functions as a neurotransmitter.
Single nucleotide polymorphisms (SNPs)	Single-base changes in a DNA sequence, for example a C to T substitution. SNPs can be found in the coding (exonic) and noncoding region (e.g., untranslated region, promoter, or intron) of a gene, and coding SNPs can be synonymous (no change in resultant amino acid) or nonsynonymous (change in resultant amino acid).
Transcription	The synthesis of an RNA copy from a sequence of DNA; the first step in gene expression.
Transcription factor binding site	A consensus sequence within a gene, promoter, or regulatory region that is recognized and bound by a transcription factor that then regulates the transcription of that gene.
Variable number of tandem repeats (VNTR)	A polymorphism within a DNA sequence that comprises varying numbers of repeating units. VNTRs can be microsatellites (usually repeat units of 2–8 bp) or minisatellites (more than 8 bp). VNTRs are hypermutable and often give rise to many alleles of different lengths.

The Serotonin Transporter

The serotonin transporter (5-HTT) is important in the termination and modulation of serotonergic neurotransmission by sodium-dependent reuptake of serotonin into the presynaptic neuron, which thus terminates the synaptic actions of serotonin and recycles it into the neurotransmitter pool. The transporter is the site of action of a number of commonly used antidepressants, such as selective serotonin reuptake inhibitors (SSRIs), fluoxetine (Prozac^{TM}), sertraline (Zoloft^{TM}), and paroxetine (Paxil^{TM}), and tricyclic antidepressants. A number of psychostimulants such as cocaine, methylphenidate, and MDMA (ecstasy) may also interact with 5-HTT. For this reason, *5-HTT* has become one of the most widely studied genes in the field of behavior, especially in the context of psychiatric disorders such as depression, neuroticism, and suicidal behavior.

The *5-HTT* Gene

The gene encoding the 5-HTT protein in humans has been mapped to the chromosomal location 17q11.2 and consists of 14 exons spanning approximately 38 kb. Several studies have focused on screening the genomic region for variation, and a number of polymorphisms have been characterized. Currently, at least 62 SNPs with frequency data within the gene, untranslated, or promoter region of *5-HTT* are listed in the Entrez SNP database (dbSNP). Most of these SNPs have not yet been investigated in the context of complex behavior, and any role they might have in regulation of the gene is still unknown.

Functional Polymorphisms of the *5-HTT* Gene

There are some regions within *5-HTT* that have been demonstrated to affect the expression of the gene.

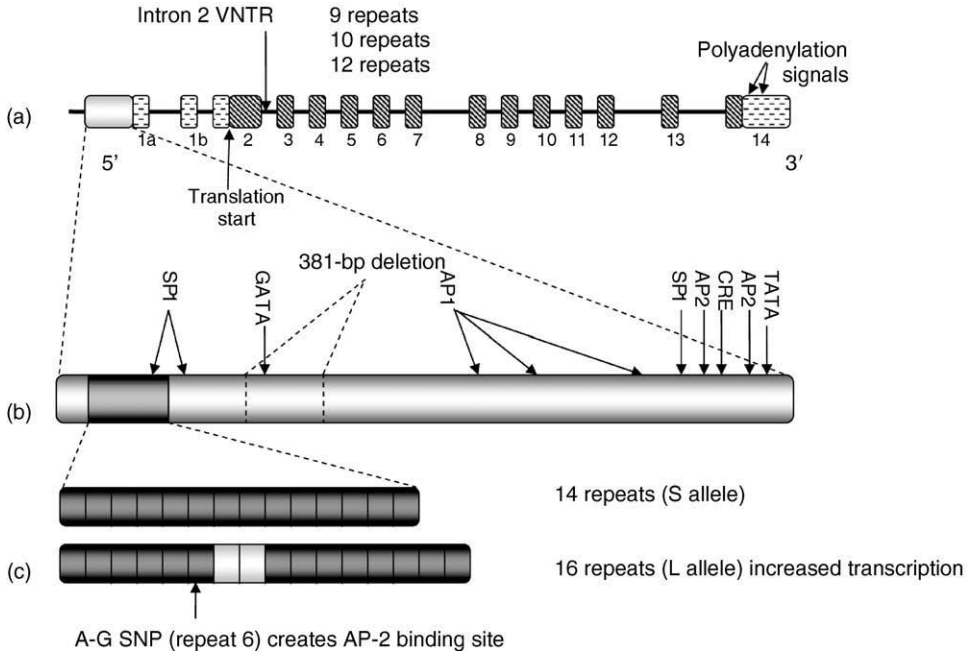


Figure 1 Serotonin transporter gene (not to scale). a, Organization of the gene; b, promoter region; c, *5-HTTLPR* as either the S or L allele. In a, the coding exons are represented by diagonally marked boxes; noncoding exons are represented by dashed boxes. The position of the promoter (light grey) and intron 2 VNTR are shown. In b, the position of some postulated transcription factor binding sites are shown by arrows; the region of the 381-bp deletion is demonstrated by dashed lines, as is the position of the *5-HTTLPR* (dark grey box). In c, the light boxes represent the position of the two repeat units that are inserted in the L allele; the position of the A-G SNP is also shown.

The most common of these are described next (see Figure 1).

The Serotonin Transporter-Linked Polymorphic Region

One of the most documented variants within *5-HTT* is the serotonin transporter-linked polymorphic region (*5-HTTLPR*; Figure 1c), a polymorphic GC-rich repetitive element within the proximal 5′ regulatory region. In humans (and indeed in primates), the *5-HTTLPR* is polymorphic, consisting of a variable number of 20- to 23-bp repeats. The *5-HTTLPR* exists as two common alleles: the short allele (S, 14 repeats) and the long allele (L, 16 repeats). There is evidence that this polymorphic region has functional significance; in JAR (human carcinoma) cells, the L allele confers higher activity than the S allele, whereas lymphoblastoid cells with the LL genotype have 1.4–1.7 times the amount of mRNA as SS or SL cells. Furthermore, the uptake of labeled serotonin in LL genotype cells is 1.9–2.2 times that in cells carrying one or two copies of the S variant. More recently, it has been shown that the *5-HTTLPR* gene actually exhibits a number of alleles of varying repeat numbers. Furthermore, the S and L alleles are themselves variable, so that, although their length is similar, SNPs occur within their repeating units. In total, at least 14 alleles of varying length and sequence have been identified, four different S alleles (14-A to 14-D), six different L alleles (16-A to 16-F), and alleles consisting of 13, 15, 18, 19, 20, and 22 repeats. The longer alleles appear to exhibit a silencer effect on the gene, having lower activity than 14- and 16-repeat variants; however, these alleles are extremely rare and appear to be found in specific populations only, such as the 20-repeat allele in the Japanese population. Recently, it has been demonstrated that an A-G SNP (dbSNP reference number rs25531) within the sixth repeat of the *5-HTTLPR*, represented by the 14-D and 16-D alleles, creates a binding site for the transcription factor AP-2 that might affect the rate of transcription of the gene.

Recently however, the functional effect of the *5-HTTLPR* polymorphism has been called into question because *in vivo* experiments failed to find the same genotypic effect as that observed in cell lines. For example, the *5-HTTLPR* genotype has no effect on the 5-HTT binding of a serotonergic ligand, a measure of 5-HTT availability, in human brain. In other words, if the *5-HTTLPR* is driving differential

transcription, it has no eventual effect on *5-HTT* availability. Similarly, some experiments suggest that 5-HTT binding in the brain does not differ by *5-HTTLPR* genotype. If these results are robust, then it may be that the expression patterns in cell-based assays do not accurately reflect the pattern of expression *in vivo*. Indeed, it has been shown that the regulation of the 5-HTT depends on the presence of the substrate (serotonin, 5-HT) rather than on any genetically driven effect. This might be an indication of the many levels of complex regulation that genes are subject to *in vivo* that may not be accurately modeled by cell-line *in vitro* approaches; the functional effect of the *5-HTTLPR* observed in cells might be adapted by the complex intra- and extra-cellular factors that brain neurons are exposed to.

Intron 2 Variable Number of Tandem Repeats

Along with the *5-HTTLPR*, there are other proposed transcriptional regulatory regions within the *5-HTT*. One example is a VNTR within intron 2 that has been shown to have differential enhancer activities in the mouse embryo and in human embryonic stem (ES) cells (**Figure 1a**). The VNTR consists of 9–12 copies of a 16- or 17-bp repeat, the 12-repeat allele (STin12) having greater enhancer activity than the 10-repeat allele (STin10) in both model systems. This ability to regulate expression is cell-specific in humans, in that it is observed in ES cells but not HeLa cells. However, as in the case of the *5-HTTLPR,* the regulation is much more complex than the effect of repeat length; the VNTR does not comprise a succession of perfect repeats, rather there are seven varying repeat sequences and the differences between alleles is described by the deletion of specific repeat elements. The functional effect of the VNTR is regulated by at least two transcription factors – CCTC-binding factor (CTCF) and Y box binding protein (YB-1) – that bind to a sequence that is deleted in the nine-repeat allele (STin9). YB-1 is inhibitory, so binding to the VNTR results in decreased expression. However, the addition of YB-1 results in a release from inhibition of the STin9 VNTR in COS-7 cells because the inhibitory site is deleted. This effect is antagonized by CTCF and is not observed with either the STin10 or STin12.

This evidence suggests that *5-HTT* expression is differentially regulated by these two regions. However, little is known about their combined effect on transcription. The only evidence to date indicates that the expression of the gene is highest in lymphoblastoid cells having both *5-HTTLPR* L/L and STin12/STin12 genotypes. The expression is lowest in cells with at least one S and one STin10 allele. Based on previous evidence of higher expression of 5-HTT with LL- and STin12/STin12-compared to S- and STin10-containing genotypes, this finding could be said to fit an *a priori* hypothesis; but some replication is necessary before any firm conclusions on the synergistic action of the *5-HTTLPR* and intron 2 VNTR can be made.

Other Regulatory Elements

The transcriptional effect of the *5-HTTLPR* and intron 2 VNTR is further compounded by the observation of splicing variation within the 5′ untranslated region, so that two alternative forms of *5-HTT* mRNA exist. Within the human brain and placenta, exon 1B is alternatively spliced and appears in transcripts to a varying degree, whereas exon 1A seems to be the principal transcription initiation point. This type of alternative splicing could help to specify cell-specific activity of the 5-HTT and ensure appropriate levels of neurotransmission in different tissues. This tissue-specific splicing of 5′ untranslated exons is also observed in the brain and gut in rats, indicating that it is a common mechanism of regulating protein expression. Furthermore, tissue-specific mRNA splicing might also be precipitated through multiple polyadenylation sites found in exon 14. This mechanism dictates alternate transcription stop sites that give rise to mRNA transcripts of varying length that, like alternative splicing of exons, might regulate tissue-specific expression. It has also been proposed that the structure of the *5-HTTLPR* is likely to result in a 381-bp deletion in the promoter region (**Figure 1b**) and that this deletion is present in 20–60% of genomic DNA derived from human brain and blood. Specific breakpoints suggest that the event is similar to V(D)J rearrangement often observed in antibody genes. This rearrangement results in looping out the 381-bp DNA sequence so that the two breakpoints are adjacent and are subsequently joined together, resulting in the deletion of the looped-out sequence.

The consequences of this deletion are not known, but it is suggested that the event is likely to be regulated by tissue- and *5-HTTLPR*-dependent mechanisms. However, this effect of mosaicism has yet to be consistently observed in multiple populations.

The Serotonin Transporter-Linked Polymorphic Region and Stress

The 5-HTT has been implicated in the stress response because serotonin concentrations are increased within the synapse when under stress. Given that the *5-HTTLPR*, through gene-regulatory mechanisms, might affect the efficiency by which serotonin is

transported back in to presynaptic neuron, the *5-HTTLPR* is an ideal polymorphism candidate for studies of stress and its outcomes.

Animal studies have proven invaluable in this context because it is possible to accurately control both psychosocial and physiological effectors, unlike in humans. In Rhesus monkeys, there are variations in response to early rearing experience that are dependent on the primate *5-HTTLPR* genotype. Monkeys that possess an S allele have significantly lower cerebrospinal fluid (CSF) 5-hydroxyindoleacetic acid (5-HIAA, a serotonin metabolite) concentrations than their homozygous L-allele counterparts. This suggests that serotonin reuptake is lowered in monkeys with an S allele, but the effect is specific to those who were peer-reared (early stressful environment) rather than parent-reared. No differential effect of genotype is seen between parent-reared monkeys. This suggests that early experience has some effect on the functioning of the serotonin system within primates and that the *5-HTTLPR* genotype becomes important only when stressful events are experienced. There also exists a knockout mouse for the serotonin transporter, a genetically manipulated mouse that is deficient in the *5-Htt* gene. Stressed mice that are heterozygous or homozygous for the knockout mutation (i.e., mice that have lower or negligible levels of 5-HTT) have a greater response to stress (measured as increased levels of stress hormone adrenocorticotropic hormone, ACTH) than homozygous wild-type mice. The levels of ACTH are the same in knockout and wild-type mice when they are not stressed, again suggesting that it is the variability in 5-HTT density that participates in interindividual differences in stress response.

In humans, neuroimaging research suggests that individual differences in stress perception are mediated by genetic variations in the *5-HTTLPR*. Individuals with one or two copies of the S allele of the *5-HTTLPR* exhibit greater amygdala neuronal activity to fearful stimuli than individuals homozygous for the L allele. This suggests that the SS- and SL-genotype individuals respond to the same stressful event to a greater magnitude than the LL individuals. In addition, it has been demonstrated that there is a differential response to brain-derived neurotrophic factor (BDNF) by *5-HTTLPR* genotype. BDNF is involved in activating the stress response, so this suggests that variability in *5-HTT* may influence variability in other downstream molecular processes that 5-HTT is associated with. Some of the strongest evidence to date that the *5-HTTLPR* can modulate the human stress response comes from a longitudinal study of almost 1000 individuals from Dunedin, New Zealand, in which it was shown that variation in the *5-HTTLPR* can modify the effect of stress on the outcome of depression. Stressful life events, such as the breakdown of a marriage or unemployment can lead to depression; however, this is not true for all individuals who experience stress. It was found that individuals who had the LL genotype of the *5-HTTLPR* were protected against the effects of stress in that their rates of depression were much lower than SS- and SL-genotype individuals with similar levels of stressful life events. This effect was seen whether the stress was experienced proximally (within the previous 5 years) or earlier (in childhood). The rates of depression in individuals who had not experienced stress were similar regardless of genotype. Consequently, evidence from these diverse sources suggests that variation in the *5-HTTLPR* can influence psychopathological reactions to stress and that it can account for some of the variability in individual response.

See Also the Following Articles

Gender and Stress; Genetic Factors and Stress; Genetic Polymorphisms in Stress Response; Social Status and Stress.

Further Reading

Benjamin, A. B., Ebstein, R. P. and Belmaker, R. H. (eds.) (2002). *Molecular genetics and the human personality.* Washington, DC: American Psychiatric Publishers.

Bennett, A. J., Lesch, K. P., Heils, A., et al. (2002). Early experience and serotonin transporter gene variation interact to influence primate CNS function. *Molecular Psychiatry* **7**, 118–122.

Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301**(5631), 386–389.

Hariri, A. R., Mattay, V. S., Tessitore, A., et al. (2002). Serotonin transporter genetic variation and the response of the human amygdala. *Science* **297**(5580), 400–403.

Heils, A., Teufel, A., Petri, S., et al. (1996). Allelic variation of human serotonin transporter gene expression. *Journal of Neurochemistry* **66**(6), 2621–2624.

Klenova, E., Scott, A. C., Roberts, J., et al. (2004). YB-1 and CTCF differentially regulate the 5-HTT polymorphic intron 2 enhancer which predisposes to a variety of neurological disorders. *Journal of Neuroscience* **24**(26), 5966–5973.

Kraft, J. B., Slager, S. L. and McGrath, P. J. (2005). Sequence analysis of the serotonin transporter and associations with antidepressant response. *Biological Psychiatry* **58**(5), 374–381.

Lesch, K. P. (1998). Serotonin transporter and psychiatric disorders: listening to the gene. *Neuroscientist* **4**(1), 25–34.

Lesch, K. P. (2001). Serotonin transporter: from genomics and knockouts to behavioral traits and psychiatric disorders. In: Briley, M. & Sulser, F. (eds.) *Molecular*

genetics of mental disorders, pp. 221–267. London: Martin Dunitz.
Murphy, D. L., Li, Q., Engel, S., et al. (2001). Genetic perspectives on the serotonin transporter. *Brain Research Bulletin* **56**(5), 487–494.
Nakamura, M., Ueno, S., Sano, A., et al. (2000). The human serotonin transporter gene linked polymorphism (5-HTTLPR) shows ten novel allelic variants. *Molecular Psychiatry* **5**(1), 32–38.

Relevant Website

Entrez SNP database. http://www.ncbi.nlm.nih.gov/SNP/

Sex Differences in Human Stress Response

B M Kudielka
University of Trier, Trier, Germany
D H Hellhammer
University of Trier, Trier, Germany
C Kirschbaum
Technical University of Dresden, Dresden, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B M Kudielka, D H Hellhammer, C Kirschbaum, volume 3, pp 424–428, © 2000, Elsevier Inc.

Sex Differences in Health and Disease
Sex Differences in Subjective Aspects of the Human Stress Response
Sex Differences in Physiological Aspects of the Human Stress Response

Glossary

Cortisol — A glucocorticoid hormone produced by the adrenal cortex. Cortisol exhibits a pronounced circadian rhythm and responds to a wide range of stressors. It has many physiological effects and virtually every single nucleated cell in the body has cortisol receptors. It is predominantly (90–95%) bound to binding proteins in blood; only 5–10% of the total plasma cortisol circulates as biologically active, unbound, free cortisol.

Hypothalamic-pituitary-adrenal (HPA) axis — A central regulatory and control system of the organism that connects the central nervous system with the endocrine system. Under stress, the hypothalamus secretes corticotropin releasing hormone (CRH), which provokes the release of adrenocorticotropic hormone (ACTH) from the pituitary. ACTH triggers cortisol secretion from the adrenal cortex. The functioning of the axis is controlled by several negative feedback loops.

Lymphocytes — White blood cells responsible for the specific responses of the immune system to antigens. Based on the tissue where the stem cells mature to the specific lymphocyte, they are called B cells or T cells. B cells mature in the bone marrow and form the cellular basis for the humoral immunity (e.g., antibody production), whereas thymus-derived T lymphocytes (e.g., T helper cells and T suppressor cells) represent the effector cells of the specific cellular immunity.

Sex steroids — Hormones produced in the gonads and, to a lesser degree, in the adrenal cortex. Estradiol and testosterone play major roles in the female and male reproductive system. In women, sex steroid concentrations show typical variations during the menstrual cycle. Sex steroids also influence the peripheral and central nervous system.

Sympathetic-adrenal-medullary (SAM) axis — Sympathetic nerves, originating in the central nervous system, stimulate the adrenal medulla, which in turn secretes epinephrine and norepinephrine. These catecholamines mobilize energy-producing mechanisms quickly and downregulate less important organ functions.

Sex Differences in Health and Disease

Men and women are at differential risks for a number of illnesses. This observation is reflected in the sex-specific prevalence rates for several diseases. Women suffer more often from autoimmune illnesses such as lupus erythematosus, multiple sclerosis, and neurodermatitis, whereas men are more prone to develop coronary heart diseases or infectious diseases. Concerning psychiatric disorders, women more often develop anxiety, depression, phobias, panic disorders,

or obsessive-compulsive disorder, whereas men more often are prone to antisocial behavior, substance abuse, or suicide. Although women consistently appear to report more physical and somatoform symptoms than men, the life expectancy for women is approximately 7 years longer than it is for men.

In addition to the clear sexual dimorphism with respect to certain diseases, sex differences exist in psychological and biological aspects of the human stress response. Investigation into sex-specific stress responses may help to reveal mechanisms that underlie somatic, psychosomatic, and psychiatric illnesses. Data showing that several diseases are associated with altered stress responses support this assumption.

On a behavioral level, psychiatric disorders such as depression and anxiety are interrelated with increased symptom reporting. On a physiological level, several of these disorders are accompanied by dysfunctional endocrine regulations. Major depression, susceptibility to infectious diseases, and cardiovascular problems seem to be associated with a hyperactive HPA axis, whereas autoimmune processes are associated more often with a hyporeactive HPA axis. The observation that (premenopausal) women (in contrast to men) seem to be protected from coronary heart disease parallels the finding that men show higher blood pressure changes and catecholamine increases under psychological stress. In addition to the described sex dimorphism with respect to disease and sex-specific stress responses in humans, it is generally accepted that exposure to stress can cause or intensify numerous diseases.

Thus, differences in health status may, at least in part, be related to sex differences in stress responses. Sex-specific stress responses do not seem to be epiphenomena of sex-specific prevalence rates of several disorders but, instead, appear to reflect causal links.

Sex Differences in Subjective Aspects of the Human Stress Response

Following Lazarus and Folkman, the extent of experienced stress is not determined by a situation or an event itself but is, instead, mediated by a threephasic appraisal process. The importance or threat of the characteristics of a situation or an event is assessed in a primary appraisal loop. Simultaneously, in a secondary appraisal the amount of available coping capacity or resources is estimated. A final reappraisal determines the extent of experienced stress.

Several studies, ranging from everyday hassles to posttraumatic stress reactions, conclude that women subjectively experience more stress than men. A sociocultural approach focusing on external environmental factors emphasizes that women's lives are characterized more often by multiple social roles and lower socioeconomic status, which are both positively correlated with stressful life events. In addition to the finding that women experience more stress in everyday life, stress vulnerability and worry disposition also appear to be higher in women compared to men. This leads to the assumption that sex differences exist in people's readiness to perceive and appraise situations as stressful. This suggestion is supported by the finding that self-descriptions concerning self-esteem or self-confidence, self-efficacy, and sense of competence are slightly but consistently lower in women than in men. Although men repeatedly showed a more favorable locus of control and attribution style in questionnaire-derived studies, Miller and Kirsch concluded that data about sex differences in attributions, locus of control, and perceptions of control during stressful episodes remain inconsistent. Recent findings by Schulz and co-workers revealed that women's higher average scores in perceived stress, chance control orientation, stress-susceptibility, depressivity, and bodily complaints, as well as their lower scores in self-esteem and self-concept of ability, may be attributed to sex differences in worry disposition.

A higher sensitivity and attention to nonspecific stress symptoms appear to bring about more episodes of perceived stress in women. In contrast to women, men seem to be less sensitive in their perception of subtle bodily symptoms, and they are less prone to appraise bodily or emotional changes as significant. Concerning physical and somatoform symptoms, higher symptom reporting in women is described as a general phenomenon rather than one restricted to certain types of symptoms. This effect persists even after adjusting for psychiatric comorbidity such as depression and anxiety, disorders that are more prevalent in women and associated with increased reports of physical symptoms. The suggestion that women are more perceptive about emotions is supported by the finding that women outperform men at all ages in the encoding and decoding of emotions.

The differences in coping styles between men and women are small. Whereas women appear to prefer emotion-focused solutions including the greater use of social support and supportive networks, men focus more often on instrumental or problem-oriented coping. Reports on sex-specific effectiveness of coping strategies contradict the former notion that women's coping strategies could be inadequate or less effective for alleviating distress.

Investigations of perceived stressfulness after confrontation with various standardized laboratory

stressors repeatedly led to similar reports of feelings of distress in men and women. Whereas performance in laboratory stress tests was not associated with sex, women tend to judge their own efforts and performances as poorer than men do.

The assessment of subjective aspects of the human stress response is always based on self-reports measured by standardized questionnaires, interviews (structured or half-structured), and open (unstructured) reports. Therefore, it should be considered that sex-related response differences to stress could also reflect sex-specific biases in the willingness (or capability) to report distress.

Sex Differences in Physiological Aspects of the Human Stress Response

Different regulatory systems of the human body are activated by stress. Endocrine stress-responsive systems such as the HPA and SAM axes help the organism adapt to increased demands and maintain homeostasis after challenge. In addition to hormonal responses, stress also provokes changes in cardiovascular and immunological parameters.

Hypothalamic-Pituitary-Adrenal Axis

Investigation of HPA axis responses to different types of stressors revealed that physical exercise (>70% VO_{2max}) and psychological challenges are potent activators of the HPA axis. Whereas the majority of studies did not reveal any sex differences in response to physical exercise, confrontation with psychosocial stimuli leads to higher adrenocorticotropic hormone (ACTH) and free cortisol levels in adult men than in women. This sex difference is conserved after menopause and in elderly individuals. Total cortisol secretion, however, does not differ between sexes in young adults. Although men and women show similar total cortisol responses after stress, a sexual dimorphism exists in the amount of bioavailable free cortisol, an observation calling for simultaneous measurements of total and free cortisol levels. These findings can, at least in part, be explained by greater ACTH pulses in men as well as by the higher sensitivity to ACTH of the adrenal cortex in women.

Observed sex differences in HPA axis stress responses may be due to sexual dimorphisms in brain functions and circulating sex steroid and corticosteroid-binding globulin (CGB) levels. Prime candidates for explaining sex differences in HPA axis stress responsiveness are the gonadal steroids. Among them, estradiol seems to exert modulating effects on HPA axis functioning, including HPA axis responsiveness and sensitivity to glucocorticoid negative feedback. Animal studies consistently revealed that estradiol has a strong stimulatory influence on HPA axis functioning with modulatory effects on mineralocorticoid and glucocorticoid receptors.

Although human data are more contradictory than evidence from animal literature, gonadal steroids also seem to exert important modulatory effects on HPA axis functioning in humans. In a series of studies from our laboratory, free cortisol responses systematically varied as a function of menstrual-cycle phase. Women in the luteal phase show free cortisol stress responses comparable to men, whereas women in the follicular phase and women taking oral contraceptives have significantly lower free cortisol responses. Although the total plasma cortisol net stress responses differ neither between men and women nor between different menstrual-cycle phases, the absolute levels are elevated in women with oral contraceptive use. A short-term estradiol treatment potentiated HPA axis responses to acute psychosocial stress in young men. Furthermore, a 2-week estradiol substitution seemed to diminish the endocrine response of postmenopausal women to exogenous corticotropin releasing hormone (CRH) after dexamethasone premedication, as observed in women with and without short-term estradiol treatment. In addition, a comparison with a third group of premenopausal women indicates that the feedback sensitivity of elderly estradiol-treated women approximates the endocrine response of premenopausal (untreated) women. However, this treatment did not change the acute HPA axis stress responses to psychosocial stress (Trier Social Stress Test). In sum, it should be acknowledged that there is still a paucity of data on the effects of gonadal steroids on the regulation of the human HPA (and SAM) axis. There is a definite need for future in-depth studies on these issues.

Sympathetic-Adrenal-Medullary Axis

Studies investigating SAM axis responses to stress showed that stress-induced epinephrine increases are higher in men than in women, suggesting that men are more sensitive to sympathetic stimulation via emotional arousal. In contrast, no sex effects are reported for norepinephrine concentrations. However, contradictory results also exist.

Reproductive hormones appear to have a significant impact on catecholamine stress responses. This assumption is confirmed first by findings of altered catecholamine levels throughout the menstrual cycle with enhanced epinephrine levels during the luteal phase. Second, comparisons between pre- and postmenopausal women revealed that stress-related epinephrine responses are increased in postmenopausal women

and that estradiol treatment reduced catecholamine reactivity to stress. Third, short-time estradiol administration modulates catecholamine reactivity in healthy young men, although the direction of the effect remains contradictory. Although the majority of studies point to an estradiol-mediated reduction in SAM axis responses to laboratory stressors, a few carefully conducted studies show sex steroids to have no or a stimulating effect.

Cardiovascular Responses

In a meta-analysis, Stoney and coworkers summarized sex differences in stress-related heart-rate and blood pressure changes. They concluded that women have higher heart rates at rest as well as slightly larger increases during challenge, whereas men exhibit higher systolic blood pressure changes. Allen and colleagues also showed higher heart-rate increases in women and higher blood pressure changes in men after confrontation with various laboratory stress tasks, supporting the conclusion that women are more likely to be cardiac reactors and men are more prone to be vascular reactors. We aggregated heart-rate responses across five different studies in healthy children and in younger and elderly adults. Accordingly, enhanced heart-rate stress reactivity can be observed in girls and young adult females compared to boys and young adult men. However, in older men and women peak heart-rate responses are comparable, with only men returning to prestressor baselines during recovery.

Although the data are inconsistent, reports of blood pressure changes as a function of menstrual-cycle phase point to sex steroids having an impact on cardiovascular functioning. The fact that coronary heart disease morbidity and mortality rates increase in women after menopause, a time characterized by the marked decrease of reproductive hormones, confirms the protective role of sex steroids. In line with this idea, reports show that postmenopausal women have a higher cardiovascular responsiveness than premenopausal women.

Whereas postmenopausal estrogen use predominantly reduces the risk of coronary heart disease in women, exogenous estrogen use in clinical trials or hormonal replacement therapy studies do not consistently confirm the protective role of estrogens. The apparently conflicting results of exogenous estrogen use on cardiovascular functioning could be explained by a dual action of estrogens. Estrogens may increase the risk of thrombo-embolic events due to an adverse influence on coagulation, and simultaneously they may help to prevent atherogenic events through favorably changed lipid and lipoprotein levels.

Immunological Responses

In general, immunological activity appears to be higher in women than in men. Whereas levels of circulating immunoglobulins are higher and antibody responses to a variety of pathogens are larger in females, sex differences in cell-mediated immunity can vary depending on the immunological parameter studied. Sex differences in basal immunological functioning may result from the stimulation of lymphoid tissues by sex steroids. In general, estrogens appear to suppress T-cell-dependent immune function, but can increase antibody reactivity to antigens. Testosterone (and other androgens) has an immunosuppressive effect and increases susceptibility to infection, although it may help to prevent the development of autoimmune processes. In sum, sex hormones have been shown to modulate a large variety of mechanisms involved in the immune response. Evidence is accumulating that shows that estrogenic and androgenic actions on cells can be mediated not only through classical intracellular receptors but also through membrane receptors on cell surfaces of compounds of the immune system. Wunderlich and coworkers investigated possible cross-talk of nongenomic testosterone signaling with other genotropic signaling pathways in the rat. Furthermore, in addition to direct receptor-mediated influences, indirect effects on immunological functioning and disease processes may be exerted by sex-steroid-induced changes in other biological systems that are interrelated with the regulation of the immune system. For example, the androgenic steroid hormone dehydroepiandrosterone sulfate (DHEAS) was recently shown to inhibit apoptosis in human peripheral blood lymphocytes through a mechanism independent of either androgen or estrogen receptors.

Although there is a voluminous literature showing the impact of acute laboratory stress on rapid immune cell changes, sexual dimorphisms in immune responses to stress in humans have not been extensively examined and the reported results remain contradictory. Whereas Kiecolt-Glaser and colleagues found that women are more likely to elicit immunological changes after controversial discussions with their husbands, as reflected by a summary immune change score (based on killer cell lysis, blastogenic responses to two mitogens, the proliferative response to a monoclonal antibody to the T_3 receptor, percentage of macrophages, and number of T lymphocytes), Scanlan and coworkers concluded that immunological parameters appear to be more affected by psychosocial stressors (hassles and caregiving) in men than in women. Recently, Pehlivanoglu and colleagues observed a shift from cellular to humoral immunity (only) in men and showed evidence that acute stress

may cause different immunological changes across the female menstrual cycle. In addition to these sexual dimorphisms, the use of oral contraceptives (OC) can also alter stress-induced immune responses in women. Furthermore, sex and OC intake may exert differential effects on the immune system by modulating glucocorticoid sensitivity of pro-inflammatory cytokine production.

In sum, immune functions at rest and under stress differ significantly between men and women, apparently due to sex-related differences in sex steroid levels.

See Also the Following Articles

Adrenal Medulla; Autoimmunity; Coping Skills; Depression Models; Economic Factors and Stress; Gender and Stress; Heart Disease/Attack; Immune Response; Menopause and Stress; Social Status and Stress; Social Support.

Further Reading

Allen, M. T., Stoney, C. M., Owens, J. F., et al. (1993). Hemodynamic adjustments to laboratory stress: the influence of gender and personality. *Psychosomatic Medicine* **55**, 505–517.

Cleary, P. D. (1987). Gender differences in stress-related disorders. In: Barnett, R. S., Biener, L. & Baruch, G. K. (eds.) *Gender and stress*, pp. 39–72. New York: Free Press.

Kiecolt-Glaser, J. K., Malarkey, W. B., Chee, M., et al. (1993). Negative behavior during marital conflict is associated with immunological down-regulation. *Psychosomatic Medicine* **55**, 395–409.

Kirschbaum, C., Schommer, N., Federenko, I., et al. (1996). Short term estradiol treatment enhances pituitary-adrenal axis and sympathetic responses to psychosocial stress in healthy young men. *Journal of Clinical Endocrinology and Metabolism* **81**, 3639–3643.

Kroenke, K. and Spitzer, R. L. (1998). Gender differences in the reporting of physical and somatoform symptoms. *Psychosomatic Medicine* **60**, 150–155.

Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H., et al. (2004). Differential heart rate reactivity and recovery after psychosocial stress (TSST) in healthy children, younger adults, and elderly adults: the impact of age and gender. *International Journal of Behavioral Medicine* **11**, 116–121.

Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H., et al. (2004). HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: impact of age and gender. *Psychoneuroendocrinology* **29**, 83–98.

Kudielka, B. M. and Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. *Biological Psychology* **69**, 113–132.

Kudielka, B. M., Schmidt-Reinwald, A. K., Hellhammer, D. H., et al. (1999). Psychological and endocrine responses to psychosocial stress and dexamethasone/corticotropin-releasing hormone in healthy postmenopausal women and young controls: the impact of age and a two-week estradiol treatment. *Neuroendocrinology* **70**, 422–430.

Lazarus, R. S. and Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Lutzky, S. M. and Knight, B. G. (1994). Explaining gender differences in caregiver distress: the roles of emotional attentiveness and coping styles. *Psychology and Aging* **9**, 513–519.

McCruden, A. B. and Stimson, W. H. (1991). Sex hormones and immune function. In: Ader, R., Felten, D. L. & Cohen, N. (eds.) *Psychoneuroimmunology*, pp. 475–793. San Diego, CA: Academic Press.

Miller, S. M. and Kirsch, N. (1987). Sex differences in cognitive coping with stress. In: Barnett, R. S., Biener, L. & Baruch, G. K. (eds.) *Gender and stress*, pp. 278–307. New York: Free Press.

Pehlivanoglu, B., Balkanci, Z. D., Ridvanagaoglu, A. Y., et al. (2001). Impact of stress, gender and menstrual cycle on immune system: possible role of nitric oxide. *Archives of Physiology and Biochemistry* **109**, 383–387.

Polefrone, J. M. and Manuck, S. B. (1987). Gender differences in cardiovascular and neuroendocrine response to stress. In: Barnett, R. S., Biener, L. & Baruch, G. K. (eds.) *Gender and stress*, pp. 13–37. New York: Free Press.

Raison, C. L. and Miller, A. H. (2003). When not enough is too much: the role of insufficient glucocorticoid signaling in the pathophysiology of stress-related disorders. *American Journal of Psychiatry* **160**, 1554–1565.

Rohleder, N., Wolf, J. M. and Kirschbaum, C. (2003). Glucocorticoid sensitivity in humans – interindividual differences and acute stress effects. *Stress* **6**, 207–222.

Scanlan, J. M., Vitaliano, P. P., Ochs, H., et al. (1998). CD4 and CD8 counts are associated with interactions of gender and psychosocial stress. *Psychosomatic Medicine* **60**, 644–653.

Schulz, P., Schlotz, W., Wolf, J., et al. (2002). [Gender differences in stress-related variables: the influence of worry-disposition]. *Zeitschrift für Differentielle und Diagnostische Psychologie* **23**, 305–326.

Stoney, C. M., Davis, M. C. and Matthews, K. A. (1987). Sex differences in physiological responses to stress and in coronary heart disease: a causal link? *Psychophysiology* **24**, 127–131.

Takahashi, H., Nakajima, A. and Sekihara, H. (2004). Dehydroepiandrosterone (DHEA) and its sulfate (DHEAS) inhibit the apoptosis in human peripheral blood lymphocytes. *Journal of Steroid Biochemistry and Molecular Biology* **88**, 261–264.

Whitacre, C. C. (2001). Sex differences in autoimmune disease. *Nature Immunology* **2**, 777–780.

Wunderlich, F., Benten, W. P., Lieberherr, M., et al. (2002). Testosterone signaling in T cells and macrophages. *Steroids* **67**, 535–538.

Young, E. A. (1995). The role of gonadal steroids in hypothalamic-pituitary-adrenal axis regulation. *Critical Reviews in Neurobiology* **9**, 371–381.

Sex Steroids, Response to Stress and Susceptibility to Depression

M V Seeman and L E Ross
University of Toronto, Toronto, Ontario, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M V Seeman, volume 3, pp 429–434, © 2000, Elsevier Inc.

Introduction
Gender Differences in Depression and Stress Syndromes
Sex Steroids
Stress
Conclusion

Glossary

Circadian rhythms	Twenty-four-hour sleep–wake cycles.
Estrous cycles	Cyclic sexual receptivity in female animals that is accompanied by specific hormonal periodicity and that lays the groundwork for impregnation.
Genomic effects	Direct effects on DNA in cell nuclei.
Growth hormone	A polypeptide hormone secreted by the anterior lobe of the pituitary gland that promotes tissue and skeletal growth and influences the metabolism of proteins, carbohydrates, and lipids.
Hypothalamic-pituitary-adrenal axis	The hormonal circuit from the hypothalamus to the anterior pituitary gland to the adrenal gland and back to the hypothalamus.
Hypothalamic-pituitary-gonadal axis	The hormonal circuit from the hypothalamus to the anterior pituitary gland to the ovaries or testes and back to the hypothalamus.
Hypothalamic-pituitary-thyroid axis	The hormonal circuit from the hypothalamus to the anterior pituitary gland to the thyroid gland and back to the hypothalamus.
Melatonin	A hormone that is secreted by the pineal gland and that regulates the onset and timing of sleep.
Neurotransmitters	Chemical messengers that cross the gap between neurons to either excite or inhibit the neighboring neuron.
Posttraumatic stress syndrome	A set of psychiatric symptoms that develop as a reaction to a traumatic event.
Prolactin	A hormone secreted by the anterior pituitary gland whose function, among others, is to stimulate the growth of milk-secreting glands in the breast in preparation for, or maintenance of, lactation.
Psychophysiological disorders	Disorders such as chronic fatigue syndrome, irritable bowel syndrome, and fibromyalgia, which are stress-related body symptoms

Introduction

From a physiological point of view, both depression and anxiety are the product of complex alterations of the hypothalamic-pituitary-thyroid axis, of prolactin, growth hormone, and melatonin irregularity, of sleep, activity, appetite, and temperature disturbance, of hypothalamic-pituitary-gonadal axis and hypothalamic-pituitary-adrenal axis deregulation, and of multiple neurotransmitter disequilibrium. Given the gender ratio of the prevalence of depression and anxiety and the generally accepted role that stressful life events and powerful emotions play in the precipitation of mood disorders, this article focuses on the effects of ovarian hormones on the body's response to stress.

Gender Differences in Depression and Stress Syndromes

Women are more susceptible than men to all forms of depression. Even in bipolar affective (manic-depressive) illness, in which the gender prevalence is equal, the frequency of depressive episodes is higher in women. Chronic, subthreshold depression (dysthymia) and episodic major depressive disorder are at least twice as common in women as they are in men. These sex differences have been replicated in all major door-to-door epidemiological surveys and in treatment-seeking population samples from around the world, with the exception of reports from countries where women are known to have reduced access to medical services. Strikingly, this sex ratio does not emerge until adolescence, when gonadal hormones are activated. Before puberty, depressive episodes are as frequent in boys as they are in girls.

Anxiety may be a symptom of a depressive episode or it may constitute the major component of a variety

of anxiety syndromes such as posttraumatic stress disorder (PTSD), phobia, panic disorder, agoraphobia, obsessive–compulsive disorder, and generalized anxiety. With the exception of obsessive–compulsive disorder, in which the gender ratio is equal, adult women are more susceptible than men to all the preceding *Diagnostic and Statistical Manual of Mental Disorders* (DSM)-defined syndromes. As in depression, the sex differences in most forms of anxiety emerge only after adolescence.

The aftermath of stressful life events may be experienced in the psychological realm, as in PTSD, or primarily in the somatic realm, as in the various psychophysiological disorders. Both forms are more prevalent in women. When exposed to a major trauma, approximately twice as many women as men develop PTSD, and women make up about 85% of PTSD sufferers whose symptoms last for over 1 year after exposure. After the World Trade Center attack, gender differences in PTSD could be explained by differences in marital status (more common among women living alone without a permanent relationship) and education (more common among women with relatively little education and low income), but significantly more women experienced major depression and significantly more men were diagnosed with a substance abuse disorder. These differences persisted after adjustment for demographic variables.

Both depression and anxiety are more severe and more frequent during the premenstrual, postpartum, and menopausal periods of women's lives. An example of experimentally elicited menstrual phase vulnerability is the Van Goozen et al. report of a study involving 58 healthy women. Deliberately arranged anger provocation evoked most emotionality in those women who were premenstrual at the time of the experiment. Bloch et al. used an analogous design to investigate susceptibility to depression in response to exogenous estrogen and progesterone among women with a history of postpartum depression. Five of eight women with a history of postpartum depression, and none of the control women, with no psychiatric history, developed significant mood symptoms during hypogonadism induced via administration of a gonadotropin releasing-hormone agonist.

In summary, depressive and anxious states, as well as stress-provoked somatic conditions, are two to three times as common in women as in men, but there is no consistent sex difference in their prevalence prior to puberty. In women, symptom severity and frequency are associated with times in the reproductive cycle when the circulating level of ovarian hormones (estrogens and progestins) are falling.

Sex Steroids

Male–Female Differences

Gender differences in illness prevalence can result from a combination of inborn, learned, and socially imposed factors, but what distinguishes the sexes during fetal development are sex-specific hormones that exert powerful effects on the fetal brain and contribute to its sexual differentiation. While the hormonal milieu does not differ appreciably for boys and girls between birth and puberty, their brains are sexually differentiated from early gestation. The dramatic hormonal changes at puberty, superimposed on a dozen or more years of differential sex role experience, exert a profound effect on the adolescent brain. In addition to median levels of sex steroids, what also differs after puberty (and this may turn out to be the crucial difference) is that only women undergo menstrual cycles. Both sexes are subject to pulsatile changes in hormone level and circadian fluctuations, but the monthly repetitive hormonal cycles interrupted by pregnancy are unique to women. It is not until very old age, when the testes stop producing testosterone, that the hormonal environment of the male brain and that of the female brain once again approach the par that existed prior to puberty. Even then, receptor densities for the various hormone agonists and their site specificity remain gender distinct.

These biological differences may play out in the context of differential exposure to social and/or psychological stress. It has been proposed that women may experience a greater frequency of stressful life events than men. Alternatively, women may be more likely than men to interpret life events as stressful. However, a recent twin study by Kendler et al. revealed that men and women were equally likely to experience an episode of major depression associated with a major life event, although the categories of significant events differed between the sexes (i.e., interpersonal events were more relevant to women; work problems were more relevant to men).

In summary, although exposure to the impact of different experiences modulates the internal cascade of neurochemical events in which sex hormones play such a crucial part, neuroendocrine differences and their sex-specific rhythms are arguably central to male–female differences in the vulnerability to depression and anxiety. Hormones determine how the male and female brain processes the experience to which it is exposed, in particular, how it differentially responds to stressful experience and powerful emotions.

Sex Steroids and Neurotransmitter Systems

Sex steroids operate through genomic effects on neurotransmitter receptor protein or through direct membrane effects on the electrical firing rate of neurotransmitter-secreting cells. The genomic effect is relatively slow (minutes to hours); the membrane effect is rapid (microseconds to minutes).

It is well recognized that the serotonergic system is important in the maintenance of mood. The effectiveness of selective serotonin reuptake inhibitors (SSRIs) in the treatment of both depression and anxiety has highlighted the serotonin system. Estrogens exert multiple effects on serotonin function, many of which are site specific. In the amygdala, a key brain structure involved in the mediation of environmental stress, estrogens increase serotonin turnover. Estrogens increase the density of serotonin 5-hydroxytryptamine 2A (5HT2A) receptors in rat cerebral cortex and nucleus accumbens. They decrease 5HT1A mRNA in the medial amygdala, piriform cortex, and perirhinal cortex, suggesting a possible pathway through which mood and memory are modulated. Throughout the brain, estrogens increase the rate of degradation of monoamine oxidase, which, in turn, leads to increased brain availability of both serotonin and catecholamines. The interaction between estrogen and the serotonin system may have clinical implications: there is recent evidence that SSRIs are effective in reducing vasomotor symptoms associated with estrogen decline during menopause, and estrogen therapies may hold promise in management of hormone-dependent psychiatric conditions, including postpartum depression. Finally, there is some evidence that concurrent estrogen therapy may augment or accelerate response to SSRIs in women.

Estrogen-responsive and progesterone-responsive neurons use the inhibitory γ-aminobutyric acid (GABA) as a transmitter, producing an anxiolytic affect. Estrogen also induces choline acetyltransferase and potentiates excitatory amino acids such as glutamate and aspartate. The estrogen effect on the dopaminergic system is complex and phasic, causing alterations in pre- and postsynaptic dopamine receptors and differential distribution of the isoforms of the dopamine D2 receptor messenger RNA. More specific to stress and depression, estrogens block both the synthesis and the depletion of catecholamines, including norepinephrine. Norepinephrine plays a major role in the autonomic stress response. Estrogens work via endogenous opiate peptides to exert an activating influence on the autonomic sympathetic response to stress, which results in vasoconstriction, quickened heartbeat, and sweating, common symptoms of anxiety.

In summary, ovarian hormones exert multiple, complex effects on neurotransmitter systems understood to be involved in depression and anxiety syndromes.

Effects on Membrane Stability

Because of their mainly activational influences, estrogens are proconvulsants in susceptible individuals, decreasing seizure threshold and increasing the magnitude of seizure activity. In experimental animals, this appears to be a direct consequence of depolarization of cellular membranes. Progesterone, on the other hand, raises the seizure threshold, an effect opposite to that of estrogen and probably responsible for catamenial or premenstrual epilepsy, when progesterone levels drop. Seizure activity (e.g., electric shock treatment) has been known for a very long time to ameliorate depressive mood.

In summary, the net effect of progesterone is to reduce anxiety but to deepen depression. The opposite is true for estrogen: levels of circulating estrogens correlate with increased alertness in animals, reduced distractibility, and shorter reaction times. The two hormones maintain a balance or homeostasis that, when disturbed, can tip a vulnerable person in the direction of mental states that are diagnosed as depression or anxiety.

Other Variables

Estrogens increase blood flow to the brain and within the brain, which speeds metabolism and increases the general efficiency of neural networks and increases the effectiveness of drug delivery. Estrogens stimulate glucose transport and metabolism, promoting the ability of neuronal networks to perform their work.

Neuronal responsiveness to steroid stimulation results from both genetic and membrane effects at multiply distributed sites involving large neural circuitry. Network effects may differ depending on the particular configuration of a specific circuit. Hormonal effects on neural network properties appear to be state dependent. In other words, effects may differ widely depending on the light–dark cycle, the sleep cycle, the activity cycle, the glucose level, the noise level, the presence or absence of an animal of the opposite sex, the general stress level, and the specific type of stressor.

Genomic Effects

Among the many genes suspected to come under the influence of estrogens are those responsible for neuronal development, migration, growth, and survival, for dendritic spine formation and density, and for antioxidant activity, synaptic plasticity,

neurodegeneration, and apoptosis secondary to glucose deprivation or other toxic event.

Steroid hormones are important epigenetic regulators. In a tissue-specific manner, they control the dosage effects of genes. They can turn genes off and on. They bind to their own receptors, a family of DNA-binding proteins located within the cell nucleus. The human estrogen receptor (ER-α) has been characterized both at the cDNA level and at the genomic level. It is transcribed from three different promoter regions. The expression of these three ER-α mRNAs appears to be regulated in a tissue-specific and, perhaps, age-specific way.

ER-α is sparsely distributed in key cognitive regions. This made it difficult to understand why estrogen replacement therapy sometimes improved cognitive functions until the recent discovery of a new estrogen receptor (ER-β), localized in brain regions associated with learning, memory, and emotion.

Once estrogen enters the cell nucleus and binds to the ER expressed in that cell, a conformational change takes place. The receptor's DNA binding domain is exposed and attaches to specific sequences of nucleotides in the nuclear DNA, thus regulating the transcription of a target gene. Adapter proteins or other transcription factors, which are cell specific, modulate the strength of the activation of the gene. Genetic transcription is one way in which estrogens modify the structure and function of brain cells.

Membrane Receptors

Estrogen also produces nongenomic effects by binding to neuronal membrane receptors. This effect is rapid, with a latency of seconds to minutes, and disappears quickly. Estrogen effects on dendritic spines in the CA1 region of the hippocampus are a good example of a rapid response. Estrogen given to an ovariectomized rat produces an increase in the number of spines on apical dendrites of CA1 pyramidal neurons. Cyclic changes in spine density can be seen throughout the estrus cycle of rats, suggesting the formation and destruction of synapses in quick time, presumably, therefore, mediated via membrane receptors. One mechanism through which dendritic spine density is increased is neurotrophin brain-derived neurotrophic factor (BDNF). Estradiol downregulates BDNF in cultured hippocampal neurons within 24 h of exposure. This increases excitatory tone in pyramidal neurons, which may lead to as much as a twofold increase in dendritic spine density.

In summary, estrogens, by a variety of means, help to increase neuronal connections. If nerve impairment occurs, they foster repair and delay or prevent secondary atrophy.

Stress

The Stress Response

The human organism has evolved an elaborate mechanism for dealing with perceived threat. Cortisol, epinephrine, and norepinephrine are automatically released, mobilizing the necessary energy for fight or flight. This primarily involves the limbic-hypothalamic-pituitary-adrenal-gonadal-thyroid axis. The integration of external stimuli converges on the medial parvocellular division of the paraventricular nucleus of the hypothalamus (mpPVN), which secretes corticotropin-releasing hormone (CRH). This hormone binds to its receptors in the anterior pituitary, stimulating adrenocorticotropic hormone (ACTH)-secreting cells. ACTH is released into the circulation and, in the adrenal cortex, activates the synthesis and release of glucocorticoids whose many actions are mediated via specialized receptors affecting and regulating genes throughout the body. At the same time, stress leads to the modulation of the secretion of thyroid-releasing hormone and gonadal-releasing hormone from the hypothalamus with subsequent downstream effects on the thyroid and the gonads. Appetite and the need for sleep are suppressed. Growth stops. The overall result is that the organism is prepared to meet the perceived threat.

A large part of the brain is devoted to coping with stress. The amygdala is also placed on alert and activates the hypothalamic-pituitary-adrenal (HPA) axis. The prefrontal cortex is involved in assessing the danger and putting a halt to the stress response when the danger is past.

Men and women show differential HPA axis responses to psychosocial stress, with women of reproductive age showing less pronounced HPA axis responses to stress than men of the same age. However, this response varies with menstrual cycle phase and menopausal status, suggesting interactions between the HPA axis and gonadal hormones.

Termination of the Stress Response

The level and duration of glucocorticoid activity must be adequate to the organism's need to deal effectively with the perceived stressor. Glucocorticoids (helped along by excitatory amino acids) are, however, potentially harmful and may cause cell death through the promotion of apoptosis. A mechanism needs to be in place for their rapid elimination. Otherwise, what is adaptive becomes, under conditions of severe or ongoing stress, an agent of pathological change. Hippocampal neurons, particularly rich in steroid receptors, are exquisitely vulnerable to glucocorticoid-induced atrophy.

Many negative feedback loops exist within the stress activation system, which effectively puts a stop to the glucocorticoid stress response once it is no longer needed. Besides inhibition of the amygdala by the thinking brain, the prefrontal cortex, mounting concentrations of cortisol in the brain automatically shut off production of CRH. Sex steroids do the same. The rapidity of both stress activation and stress termination is affected by the circadian rhythm in plasma glucocorticoid levels, being most responsive when steroid levels are at their lowest, in the late evening in diurnal animals.

Stress Response Termination in Depression

Efficient termination of the stress response may be impaired in depression. For instance, the left prefrontal cortex has been shown in functional imaging studies to be impaired in major depression. As mentioned previously, depression and stress have long been associated. Adverse life events trigger depressive episodes, and depression itself constitutes an ongoing stressor to the organism. The experience of depression is one that precludes pleasure and hope, prevents sleep, and diminishes appetite, metabolic activity, autonomic function, neuroendocrine processes, growth, and reproduction.

There has been a long tradition of cortisol measurement in depressive states, culminating in the mid-1970s with the introduction of the Dexamethasone Suppression Test (DST) as a diagnostic test. The assumption was that suppression would be deficient in depression, and this has turned out to be true in at least some forms of depression.

This fact may explain some of the cognitive impairment that accompanies depression when glucocorticoid levels remain high and result in atrophy of hippocampal cells. Gonadal steroid receptors, with variations from one time period to another, are also expressed in the hippocampus. In animal experiments, androgens have been shown to downregulate glucocorticoid receptors in the hippocampus. Estrogens have also been reported to protect against glucocorticoid-induced hippocampal atrophy. Although both androgens and estrogens are potentially protective, both female rats and female monkeys undergoing various forms of experimentally induced stress fare better than their male counterparts with respect to secondary hippocampal damage.

In summary, a sustained stress response produces hippocampal degeneration and atrophy of apical dendrites of hippocampal pyramidal neurons. The interaction among glucocorticoids, neuronal growth factors, neurotransmitters, and sex steroids is complex and still in the process of clarification. Acute responses to stress, when prolonged, cause damage.

Conclusion

It would appear from animal work that women might have an advantage over men with respect to neuronal protection in the face of stress. Why, then, are they more prone than men to depression and anxiety?

The integrity of circadian rhythms and sleep–wake cycles seems to play a part in depression. People with depression feel worse in the morning than in the evening. They cannot sleep. They wake early. Resynchronizing the sleep cycle is an effective treatment for depression. The stress response cascade is modulated by circadian rhythm, and estrogens appear to be important circadian pacemakers. Is it possible that the monthly periodicity of ovarian hormones destabilizes these regular daily rhythms and thus contributes to women's special vulnerability to depression? The monthly cycling may sensitize target receptors and may in fact kindle these mechanisms in such a way that high circulating levels of ovarian hormones are needed to maintain equilibrium. This means that vulnerability to stress-related episodic depression and anxiety is heightened during periods of relative hormone decline.

That being said, it is also important to keep in mind that the physiological stress response cascade and its sequelae may be triggered by emotions other than fear of perceived threat. It may be precipitated by strong feelings of many sorts, such as frustration, empathy, or guilt, for instance, and these feelings may be relatively ubiquitous in women, prone as women are to relatively close interpersonal attachments and larger social networks when compared to men. Nurturant proclivities may be mediated by prolactin, produced in the anterior pituitary. The activities of this hormone (more active in women) may help to explain the sex ratio of adult depressive disorders.

In summary, women are more prone than men to suffer from anxiety and depression in their many forms. The mechanisms involved may reflect a tendency to hyperarousal or prolonged arousal when responding to a stressor. The sex differences may be due to relative inefficiency in turning off the stress signal, perhaps resulting from monthly kindling by fluctuations in ovarian hormones or, alternatively, from a relatively wider range of stimuli eliciting the stress response in women.

See Also the Following Articles

Anxiety; Depression and Manic-Depressive Illness; Estrogen; Gender and Stress; Sex Differences in Human Stress Response.

Further Reading

Bloch, M., Schmidt, P. J., Danaceau, M., et al. (2000). Effects of gonadal steroids in women with a history of postpartum depression. *American Journal of Psychiatry* **157**, 924–930.

Bora, S. H., Liu, Z., Kecojevic, A., Merchenthaler, I. and Koliatsos, V. E. (2005). Direct, complex effects of estrogens on basal forebrain cholinergic neurons. *Experimental Neurology* **194**, 506–522.

Dalla, C., Antoniou, K., Drossopoulou, G., et al. (2005). Chronic mild stress impact: are females more vulnerable? *Neuroscience* **135**, 703–714.

Davydov, D. M., Shapiro, D., Goldstein, I. B. and Chicz-DeMet, A. (2005). Moods in everyday situations: effects of menstrual cycle, work, and stress hormones. *Journal of Psychosomatic Research* **58**, 343–349.

Fitch, R. H. and Denenberg, V. H. (1998). A role for ovarian hormones in sexual differentiation of the brain. *Behavioral and Brain Sciences* **21**, 311–352.

Holbrook, T. L., Hoyt, D. B., Stein, M. B. and Sieber, W. J. (2002). Gender differences in long-term posttraumatic stress disorder outcomes after major trauma: women are at higher risk of adverse outcomes than men. *Journal of Trauma* **53**, 882–888.

Huttner, R. P. and Shepherd, J. E. (2003). Gonadal steroid, selective serotonin reuptake inhibitors, and mood disorders in women. *Medical Clinics of North America* **87**, 1065–1076.

Kajantie, E. and Phillips, D. I. W. (2005). The effects of sex and hormonal status on the physiological response to acute psychosocial stress. *Psychoneuroendocrinology* **31**, 151–178.

Kendler, K. S., Thornton, L. M. and Prescott, C. A. (2001). Gender differences in the rates of exposure to stressful life events and sensitivity to their depressogenic effects. *American Journal of Psychiatry* **158**, 587–593.

McEwen, B. S. and Magarinos, A. M. (2001). Stress and hippocampal plasticity: implications for the pathophysiology of affective disorders. *Human Psychopharmacology* **16**(S1), S7–S19.

Merchenthaler, I., Lane, M. V., Numan, S. and Dellovade, T. L. (2004). Distribution of estrogen receptor alpha and beta in the mouse central nervous system: *in vivo* autoradiographic and immunocytochemical analyses. *Journal of Comparative Neurology* **473**, 270–291.

Razandi, M., Pedram, A., Merchenthaler, I., Greene, G. L. and Levin, E. R. (2004). Plasma membrane estrogen receptors exist and function as dimers. *Molecular Endocrinology* **18**, 2854–2865.

Steiner, M., Dunn, E. and Born, L. (2003). Hormones and mood: from menarche to menopause and beyond. *Journal of Affective Disorders* **74**, 67–83.

VanGoozen, S. H., Frijda, N. H., Wiegant, V. M., Endert, E. and Van de Poll, N. E. (1996). The premenstrual phase and reactions to aversive events: a study of hormonal influences on emotionality. *Psychoneuroendocrinology* **21**, 479–497.

Weissman, M. M., Neria, Y., Das, A., et al. (2005). Gender differences in posttraumatic stress disorder among primary care patients after the World Trade Center attack of September 11, 2001. *Gender Medicine* **2**, 76–87.

Sex-Specific Effects of Early Social Stress in Mammals: A Study in Guinea Pigs

S Kaiser and N Sachser
University of Muenster, Münster, Germany

© 2007 Elsevier Inc. All rights reserved.

Prenatal Stress
Nonsocial versus Social Stressors
A Study in Guinea Pigs
Comparison with Other Nonhuman Mammals
Neuroendocrine Mechanisms of Prenatal Social Stress
Adaptive Function of Prenatal Social Stress

Glossary

Demasculinization	Loss or attenuation of male-typical traits.
Feminization	Enhancement of female-typical traits.
Maternal effects	Mothers' control over offspring development.
Prenatal stress	When stressors acting on pregnant females result in significant consequences for the offspring during fetal development and/or later in life.
Social environment	Part of an individual's environment that is constituted by other members of the same species.
Social stress	Stress due to members of the same species.

Prenatal Stress

Internal and external stressors constantly challenge organisms. Animals and humans respond to environmental challenge by activating a variety of coping systems (e.g., the hypothalamic-pituitary-adrenocortical [HPA] and the sympathetico-adrenomedullary [SAM] system). These systems are activated to maintain the internal milieu within a normal physiological range (i.e., homeostasis). This is crucial for normal pre- and postnatal development. The activation of the HPA axis results in an elevated secretion of glucocorticoids (e.g., cortisol, corticosterone), while the activation of the SAM axis causes an increase in serum catecholamine concentrations (e.g., epinephrine, norepinephrine). If stressors act on a pregnant female, concentrations of maternal hormones will change. Because the internal environment of the pregnant animal constitutes much of the embryo's external environment, it is reasonable to suggest that the former may affect fetal development. It is well-known that stressors acting on the organism during the pre- and/or perinatal phase can have distinct and long-term effects on behavior, reproductive functions, and the immune as well as the neuroendocrine and autonomic systems. For example, it has been reported numerous times that stress applied to pregnant rats results in demasculinization (i.e., lower levels of copulatory behavior) and feminization (i.e., higher levels of lordosis behavior, a female-typical sexual behavior) of male offspring compared to sons of nonstressed mothers.

Nonsocial versus Social Stressors

In most experimental studies on the effects of early stress, pregnant nonhuman mammals were subjected to nonsocial stressors (e.g., regimens of heat, bright light, or restraint). However, most of these stressors do not exist in the natural world of the animals and so it is questionable whether they represent ecologically relevant stimuli.

In their natural habitat, animals have to cope with a variety of social stressors that are indicative of their ecological niche. They have to adjust to stressors in the physical aspect of their environment (e.g., inclement weather) as well as in the biotic world that surrounds them (e.g., predators, food shortage). A major part of an individual's biotic environment consists of other members of the same species, which can be defined as his or her social world. In fact, a majority of human and nonhuman animals' daily expectations, motivations, and behaviors are directed to encounters with conspecifics. On the one hand, this social world can support good welfare and health (e.g., through the effects of social support). On the other hand, it can result in severe degrees of stress that can eventually lead to disease and even death (e.g., in the case of social defeat, social instability, or crowding). Thus, the social environment can represent a most effective stressor, and during pregnancy it can be crucial for the development of the offspring as well.

A Study in Guinea Pigs

In nonhuman mammals, studies on prenatal social influences on offspring development have been conducted in several species, including mice, rats, bank voles, squirrel monkeys, and pigs. The most comprehensive data so far have come from guinea pigs. In this species social stimuli have widespread neuroendocrine effects, and it is argued that this species is well suited for the study of socioendocrine effects throughout the life span and can be a valuable complement to nonhuman primate research in this area. Guinea pigs in prenatal stress research were kept in two different social situations: in a stable and an unstable social environment. In the stable social environment, the group composition (one male, five females) was kept constant; in the unstable social environment, two females from different groups were exchanged every third day. The instability of the social environment during pregnancy and lactation alters the development of the offspring in a sex-specific fashion (for a summary see Table 1).

Table 1 Summary of the effects of early social stress in guinea pigs[a]

	Females	*Males*
Behavior	SF masculinized	SM infantilized
Body weight	SF = CF	SM = CM
Serum cortisol	SF = CF	SM delayed decrease with age
Serum testosterone	SF >>> CF	SM = CM
Adrenal TH	SF >>>> CF	SM <<<< CM
Adrenal weight	SF >>> CF	SM = CM
Leucocytes	SF << CF	SM = CM
Granulocytes	SF << CF	SM = CM
ER-α ARC	SF >>>> CF	SM = CM
ER-α MPOA	SF > CF	SM = CM
ER-α CA1	SF >> CF	SM << CM
AR ARC	SF >>>> CF	SM <<<< CM
AR MPOA	SF >>>> CF	SM << CM
AR CA1	SF >> CF	SM = CM

[a] CF, adult control females; SF, early stressed females; CM, control males; SM, early stressed males; DHEA, dehydroepiandrosterone; DHEAS, dehydroepiandrosteronesulfate; adrenal TH, activities of the adrenal tyrosine hydroxylase; ER-α, estrogen receptor-α, neurons with nuclear staining; ARC, arcuate nucleus; MPOA, medial preoptic area of the hypothalamus; CA1, pyramidal cell layer of Ammon's horn of the hippocampus; AR, androgen receptor, neurons with nuclear staining. Statistics: </>, $p < 0.1$; <</>>, $p < 0.05$; <<</>>>, $p < 0.01$; <<<</>>>>, $p < 0.001$; =, $p > 0.1$ (for details see text).

Effects on the Development of the Daughters

Daughters whose mothers lived in an unstable social environment show a most conspicuous behavioral masculinization: they display significantly higher amounts of male-typical courtship and play behavior later in life than daughters whose mothers lived in a stable social environment. The behavioral masculinization of prenatally stressed adult female guinea pigs is accompanied by increased testosterone concentrations. (Even though testosterone is mainly found in male mammals, produced predominantly by the testis, it can also be found in females. In female mammals, testosterone is secreted by the ovaries as well as the adrenal cortex; however, in much lower concentrations compared to males.) Furthermore, daughters whose mothers lived in an unstable social environment show elevated activities of the SAM system compared to daughters whose mothers have lived in a stable social environment. This activation of the SAM system is indicated by greater adrenal tyrosine hydroxylase activity. This enzyme catalyzes the rate-limiting step of the catecholamine (e.g., epinephrine, norepinephrine) biosynthesis. Finally, daughters whose mothers lived in an unstable social environment show a male-typical expression pattern of androgen receptors (ARs) in parts of the limbic system, that is, significantly more neurons with nuclear staining are present in the medial preoptic area (MPOA) and the arcuate nucleus (ARC) of the hypothalamus. Furthermore, an upregulation of ARs in the CA1 region of the hippocampus and of estrogen receptor-α (ER-α) in the MPOA, ARC, and CA1 is found in daughters whose mothers lived in an unstable social environment. The MPOA plays an essential role in the control of many aspects of sexual behavior, such as masculine behavior and sexual preference; the ARC is important for the regulation of gonadotropin-releasing factors; and the hippocampus is involved in the regulation of vegetative functions as well as in learning and memory.

Effects on the Development of the Sons

Different effects are found in early stressed males. Sons whose mothers lived in an unstable social environment display behavioral patterns, that are usually shown only by very young male guinea pigs (e.g., resting with bodily contact) more frequently and until an older age than sons whose mothers lived in a stable social environment. In addition, the courtship behavior of these sons is frequently integrated into play behavior. These effects can be regarded as behavioral infantilization, which corresponds with delayed development of the adrenocortical system (i.e., the typical decrease in concentrations of cortisol, the main glucocorticoid in guinea pigs, with age is delayed) and – in contrast to daughters – decreased activity of the SAM system indicated by lower tyrosine hydroxylase activity. Finally, males whose mothers lived in an unstable social environment show a downregulation of AR expression in the MPOA and the ARC of the hypothalamus and of ER-α in the CA1 region of the hippocampus compared to males whose mothers lived in a stable social environment.

Comparison with Other Nonhuman Mammals

As these studies in guinea pigs show, prenatal factors can influence male and female offspring in different ways: the behavior of the female offspring is masculinized, whereas the behavior of the male offspring is infantilized. Interestingly, similar effects have been found in other studies investigating effects of prenatal social stress on offspring development, even though different stressors and different species were used. One common trait of prenatally stressed females can be described as a masculinization. In guinea pigs, this masculinization affects behavior, hormonal state, and sex steroid receptor distribution in specific parts of the limbic system. Aspects of masculinization also exist in mice and bank voles. A second characteristic feature of females prenatally exposed to social stress is a more-or-less severe impairment of reproductive function later in life. For example, in mice, less receptivity, delayed puberty, and altered estrus cycles are described. In male offspring, prenatal social stress frequently leads to a less pronounced expression of male-typical traits and/or a delay in development. In guinea pigs, this affects behavior, endocrine systems, and sex steroid receptor distribution in specific parts of the limbic system and is designated as infantilization. In other species (e.g., mice, rats and bank voles), similar effects are described in terms of demasculinization (i.e., the loss or attenuation of male-typical traits) and feminization (i.e., the expression of female-typical traits). Furthermore, some studies report an impairment in the function of reproductive organs, poorer motor abilities, altered emotional behavior, and effects on learning and memory.

In those few studies that consider offspring of both sexes, sex-specific or even sex-reversed effects are described. For example, in squirrel monkeys, prenatally stressed daughters show higher levels of immunoglobulin G (IgG), whereas in sons IgG concentrations decline. However, in most studies either the effects of prenatal social stress are investigated in only one sex or the sex is not identified.

Neuroendocrine Mechanisms of Prenatal Social Stress

A single neuroendocrine pathway clearly cannot explain the consequences of prenatal social stress. Rather, a variety of mechanisms is probably involved. For many aspects, prenatal social stress affects the offspring in a sex-specific and sex-reversed way. Furthermore, at least in guinea pigs, female offspring seem to be more susceptible to this stress than male offspring (see Table 1). However, a higher responsiveness to prenatal social stress in males has also been shown.

Pituitary Adrenocortical System

There is general agreement that maternal steroids play a decisive role in shaping fetal brain development. Most authors favor a role of the maternal hypothalamic pituitary adrenocortical (HPA) system in mediating the effects of the social environment during pregnancy on the offspring's behavior. Prenatal social stress often results in an activation of the maternal HPA system, and as a consequence the secretion of glucocorticoids is elevated. Thus, in male offspring mainly two different mechanisms are discussed: the high maternal glucocorticoid concentrations can influence the development of the fetus both directly and indirectly. In the male fetus a testosterone surge during a critical period of sexual differentiation is necessary for the masculinization of brain and behavior. It is known that high glucocorticoid concentrations can suppress testosterone secretion. Thus, the elevated glucocorticoid concentrations resulting from prenatal stress can eliminate or attenuate the testosterone surge in the male fetus. This can cause the behavioral demasculinization of the prenatally stressed males.

In female offspring a different mechanism is discussed. An activation of the HPA system triggers the secretion not only of glucocorticoids but also of androgens from the adrenal cortex. (The adrenal cortex consists of three different zones: the outermost zona glomerulosa [secreting mineralocorticoids like aldosterone], the middle zona fasciculata [secreting glucocorticoids], and the innermost zona reticularis [secreting androgens].) Since androgens are able to cross the placenta, they might cause a masculinization of the female embryonic hypothalamus during a critical period of sexual differentiation.

Sympathetic Adrenomedullary System

The preceding hypotheses assume that the pituitary adrenocortical system is involved during pregnancy in mediating the influence of the social environment on the development of the offspring, and numerous experiments support this view. However, there are also indications that the HPA axis does not necessarily need to be involved. At least in some cases, an alternative pathway, which may include the maternal sympathetic adrenomedullary system, seems more likely, as can be seen from studies in guinea pigs. Surprisingly, pregnant females of this species living in stable and unstable social conditions do not differ in plasma concentrations of cortisol. Even the introduction into a group of unfamiliar conspecifics does not result in increased cortisol levels 2 h later. Thus, ongoing activity of the pituitary adrenocortical system in pregnant females living under stable and unstable social conditions is probably mostly comparable. In contrast to cortisol, serum dehydroepiandrosterone (DHEA) and dehydroepiandrosterone sulfate (DHEAS) concentrations are lower in pregnant female guinea pigs living in an unstable social environment compared to controls. DHEA and its sulfate are, in addition to cortisol, the major hormones produced by the adrenal cortex. DHEAS *per se* has no effect, but it serves as a hormone pool for DHEA. Although DHEA is a weak androgen, it can exert both estrogenic and androgenic effects. This may be due to the fact that both DHEAS and DHEA are converted to testosterone and 17β-estradiol in nonendocrine cells such as macrophages.

It has been reported that pregnant female guinea pigs living in an unstable social environment show a behavioral activation. This could be the reason for the reduced androgen concentrations of such females, because a behavioral activation usually is correlated with an increased activity of the SAM system. If so, the decreased DHEAS and DHEA concentrations probably result from the elevated level of catecholamine production in the adrenal medulla, since it is well known that catecholamines suppress DHEA(S) secretion. Due to low serum levels of maternal DHEAS, lower serum levels of estrogens or androgens might be expected in the fetus. If so, a significant impact on fetal brain development might occur.

A recent experiment supports the hypothesis that there is a relationship between the decreased androgen concentrations of pregnant females living in an unstable social environment and the behavioral infantilization of their male offspring. The decreased androgen concentration is experimentally mimicked by application of the antiandrogen flutamide during the sensitive phase of fetal CNS sexual differentiation. Sons of mothers who have received the antiandrogen during pregnancy show a similar behavioral infantilization later in life, as do sons whose mothers have lived in an unstable social environment during pregnancy.

The results regarding the masculinized daughters of females living in an unstable social environment are confusing. In mammals, masculinization occurs if sufficient androgens are available during a critical period of sexual differentiation. In the case of insufficient concentrations of androgens, a female pattern develops. In guinea pigs, however, pregnant females who have lived in an unstable social environment show decreased androgen concentrations, but the daughters are behaviorally masculinized. Thus, for guinea pigs it seems unlikely that elevated maternal androgens resulting from social stress during pregnancy bring about a behavioral masculinization of their daughters. Nevertheless, androgens do seem to be involved: masculinized daughters show male-typical patterns of androgen receptor expression in the MPOA and the ARC, two brain areas that are well known to play an essential role in the control of sex-typical behaviors and the regulation of gonadotropin-releasing factors. However, at present the neuroendocrine mechanism is not clear.

Adaptive Function of Prenatal Social Stress

The characteristic traits of individuals who were exposed to environmental stimuli (stressors) during pregnancy (e.g., masculinization of daughters, feminization of sons, heightened levels of anxiety, increased cortisol responses to stressors) are often regarded as deviation from the norm or even as pathological. Alternatively, and in accordance with current evolutionary theory, these traits might represent adaptive maternal effects as well. That is, mothers try to maximize their own Darwinian fitness by efficiently adjusting their offspring to the current environmental conditions.

Effects of Prenatal Social Stress in Guinea Pigs: Pathology or Adaptation?

If prenatal influences on behavior indeed are adaptive, the following questions arise: What is the benefit of being a masculinized daughter? What is the benefit for sons who show a less pronounced expression of male-typical traits and/or a delay in development? So far, there are no empirical data to answer these questions, but reasonable hypotheses in accordance with evolutionary theories on maternal effects can be derived.

Fluctuations of natural populations of small mammals with alternating states of high and low population density are a well-known phenomenon. They also have been documented in natural populations of the feral ancestor of the domestic guinea pig. In this species, population densities can change considerably, even from year to year. Under high-density conditions, social instability is a common trait, whereas low population densities are characterized by stable social situations. Hence, pregnant female guinea pigs can experience completely different degrees of social stability not only under experimental laboratory conditions but also in their natural habitats.

If a pregnant female that lives in a high-density population has the possibility to preprogram her daughter so that she will fit optimally into a high-density situation, what should she do? A reasonable answer is to masculinize her female offspring in order to make them more robust and/or more competitive. These traits could help her daughter more efficiently defend important resources, such as food and shelter, which are scarce in high-density situations. However, a characteristic feature of females that are masculinized behaviorally due to prenatal stress is a more-or-less severe impairment of reproductive function later in life. Thus, under high-density conditions, there is a trade-off between the benefits of a behaviorally and endocrinologically masculinized phenotype and the costs of decreased reproductive success. Nevertheless, under such conditions, the chances of reproductive success would be expected to be greater in masculinized than in nonmasculinized females. Under low-density conditions, sufficient resources are available and fewer competitive abilities are needed, so it is more beneficial to invest time and energy into reproductive effort rather than to build and maintain a male-typical phenotype to defend resources. Thus, under low-density conditions, chances of reproductive success would be greater in nonmasculinized females compared to masculinized ones.

Similarly, if mothers program their male offspring prenatally such that a less pronounced expression of male-typical traits and/or a delay in development occur, they will provide them with a promising reproductive strategy under conditions of high density.

See Also the Following Articles

Adrenal Medulla; Behavior, Overview; Corticosteroids and Stress; Environmental Factors; Gender and Stress; Hypothalamic-Pituitary-Adrenal; Social Stress, Animal Models of; Steroid Hormone Receptors.

Further Reading

Birkhead, T., Schwabl, H. and Burke, T. (2000). Testosterone and maternal effects-integrating mechanisms and function. *Trends in Ecology & Evolution* **15**, 86–87.

Cushing, B. S. and Kramer, K. M. (2005). Mechanisms underlying epigenetic effects of early social experience: the role of neuropeptides and steroids. *Neuroscience and Biobehavioral Reviews* **29**, 1089–1105.

de Kloet, E. R., Sibug, R. M., Helmerhorst, F. M. and Schmidt, M. (2005). Stress, genes and the mechanism of programming the brain for later life. *Neuroscience and Biobehavioral Reviews* **29**, 271–281.

Dufty, A. M. Jr., Clobert, J. and Müller, A. P. (2002). Hormones, developmental plasticity and adaptation. *Trends in Ecology & Evolution* **17**, 190–196.

Gandelman, R. (1992). *Psychobiology of behavioral development*. New York: Oxford University Press.

Kaiser, S. and Sachser, N. (2005). The effects of prenatal social stress on behaviour: mechanisms and function. *Neuroscience and Biobehavioral Reviews* **29**, 283–294.

Von Holst, D. (1998). The concept of stress and its relevance for animal behavior. In: Lehman, D. S., Hinde, R. & Shaw, E. (eds.) *Advances of the study of behavior*, pp. 1–131. London: Academic Press.

Ward, I. L. (1972). Prenatal stress feminizes and demasculinizes the behavior of males. *Science* **175**, 82–84.

Weinstock, M. (1997). Does prenatal stress impair coping and regulation of hypothalamic-pituitary-adrenal axis? *Neuroscience and Biobehavioral Reviews* **21**, 1–10.

Sexual Assault

N C Feeny and T J Linares
Case Western Reserve University, Cleveland, OH, USA
E B Foa
University of Pennsylvania, Philadelphia, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by N C Feeny and E B Foa, volume 3, pp 435–439, © 2000, Elsevier Inc.

Prevalence of Traumatic Events and Sexual Assault
Psychological Sequelae of Sexual Assault
Treatment of Sexual Assault-Related PTSD
Predictors of PTSD
Conclusion

Glossary

Arousal	A cluster of PTSD symptoms that includes concentration problems, anger, exaggerated startle response, sleep disturbance, and being overly alert.
Avoidance	A cluster of PTSD symptoms including behavioral and psychic avoidance, such as avoidance of thoughts, feelings, and reminders of the trauma; emotional numbing; loss of interest in activities; feeling disconnected from others; psychogenic amnesia; and a sense of foreshortened future.
Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV)	A reference manual used by mental health professionals that includes descriptions and diagnostic criteria for psychiatric disorders.
Posttraumatic stress disorder (PTSD)	An anxiety disorder that develops in response to a traumatic event and is characterized by three clusters of symptoms re-experiencing, avoidance, and arousal. PTSD is diagnosed after a 1-month period of sustained symptoms following the traumatic event.
Prevalence	How widespread a disorder or an event (e.g., sexual assault) is at a given time.
Re-experiencing	A cluster of PTSD symptoms that includes experiencing the trauma in the form of nightmares, flashbacks, intrusive distressing thoughts, or becoming intensely emotionally upset or having physiological arousal upon exposure to reminders of the trauma.
Trauma	An event experienced or witnessed, in which an individual is subjected to the threat of death, injury, or physical integrity, and during which an individual feels terrified, horrified, or helpless.

Prevalence of Traumatic Events and Sexual Assault

Estimates of exposure to traumatic events vary, but on the whole, they are quite high. Several studies have examined rates of traumatic experiences in the general population. In a nationally representative sample of approximately 6000 people, ages 15–54, 60% of men and 51% of women reported experiencing at least one traumatic event in their lifetime. Similar findings emerged from a second study using a large, racially diverse sample; 74% of men and 65% of women reported experiencing at least one traumatic event. Somewhat lower prevalence was found in a sample of young adult health maintenance organization (HMO) members ($n = 1007$): 43% of men and

39% of women reported a traumatic experience in their lifetime. This discrepancy may be accounted for by differences in the samples; members of an HMO may be of higher socioeconomic status than the general population, which may have shielded them from a high-risk level of exposure to trauma. Studies examining rates of trauma exposure internationally have found similarly high levels of exposure. In a large sample of four cities in Mexico, 76% of the participants reported at least one traumatic event in their lifetime, and the majority of the participants experienced multiple versus single traumas.

Another large epidemiological study examined rate of exposure to traumatic events among a representative sample of 4000 women. Consistent with other results, almost 70% of women reported having experienced a traumatic event during their lifetime. In addition to these high rates of single trauma exposure, prevalence data have shown women to be at high risk for experiencing multiple episodes of violence throughout their lives.

Women are more likely than men to experience the traumas of rape and sexual assault. In the literature, rape is typically considered oral or vaginal penetration, while sexual assault is a more broad term that encompasses the use of pressure and coercion to force sexual contact. Lifetime prevalence estimates of rape and sexual assault among women vary widely across studies, ranging from 2 to almost 30%. Among adult women, several studies estimated prevalence of completed rape to be 20%. Others, however, have documented much lower rates, ranging from 2 to 14%.

In an epidemiological study of 4008 women, the prevalence of completed rape was 13%; other types of sexual assault were reported by 14% of the sample. Prevalence of childhood sexual abuse is similarly high. In a nationally representative sample of 2626 adults, almost one-third (27%) of the women reported a history of sexual abuse (e.g., being touched, grabbed, kissed, shown genitals) during childhood. Of these women who reported childhood sexual abuse, 49% (13% of the entire sample) reported actual or attempted intercourse. Among adolescents, one study found that 43% of the sample had experienced a traumatic event, of which rape accounted for 2% of the traumas. Further, female adolescents were significantly more likely to report a completed rape than males, which highlights young women's risk for sexual assault. International samples also report similar levels of sexual assault in women. Of the types of trauma reported by women in the Mexico study, nearly 15% reported some type of sexual assault (10.5% sexual molestation, 3.9% sexual assault). Since sexual assault is less likely to be reported to the police than other crimes, it is safe to assume that the rates mentioned above are an underestimation of the actual prevalence of sexual assault.

Psychological Sequelae of Sexual Assault

Posttraumatic Stress Disorder (PTSD)

Sexual assault is associated with a variety of psychological difficulties. Among the most common are symptoms of posttraumatic stress, a set of psychological and physical symptoms that follow a traumatic experience. In some cases, these symptoms persist beyond the immediate aftermath of the trauma and develop into posttraumatic stress disorder (PTSD), an anxiety disorder that includes symptoms of arousal, avoidance, and re-experiencing, which lasts for more than 1 month and causes significant impairment in social or occupational functioning. This disorder affects from 1 to 12% of the general population and close to 25% of those exposed to a traumatic event. Women appear to be twice as likely to develop PTSD as men (10% of women vs. 5% of men).

Retrospective and prospective studies show female victims of sexual assault to be at particular risk for developing PTSD. For example, Norris examined the impact of 10 common traumatic events (e.g., robbery, assault, disaster) in a large, racially diverse sample. Of all the traumatic events surveyed, sexual assault yielded the highest rate of PTSD (14% lifetime). Similarly, Resnick et al. found that 12% of rape victims had current PTSD compared to 3% of victims of nonviolent traumas. Using a national sample of 1500 women, rape and sexual assault led to the highest lifetime (39%) and current (13%) rates of PTSD. In a community sample of 400 adult females, of women who had experienced a completed rape, 57% met lifetime diagnostic criteria and 17% met criteria for current PTSD. Lower rates (37% lifetime and 11% current) were found in females who had experienced a nonsexual assault.

Two prospective assessment studies have examined the prevalence of PTSD following sexual or nonsexual assault. In a study of female rape victims, 94% of the women met symptom criteria for PTSD within 2 weeks following the assault. Three months after the sexual assault, rates of PTSD had declined considerably; however, 47% continued to meet PTSD criteria. Women who did not meet criteria for PTSD 3 months after the assault showed a steady reduction in symptoms over the 12 weeks of assessment. Minimal symptom improvement was seen for those women

who continued to meet PTSD criteria 3 months after the trauma. Follow-up assessments conducted 6 and 9 months after the assault revealed that 42% of women continued to meet criteria for PTSD at each time point. In a second study, Riggs et al. found that 90% of rape victims and 62% of nonsexual assault victims met symptom criteria for PTSD within 2 weeks of an assault. At the 1-month assessment, the rate was 60% for rape and 44% for nonsexual assault survivors. At 3 months, the rate had dropped to 51 and 21% for rape and nonsexual assault survivors, respectively. Similar findings have been shown in adolescents, with those who experienced a sexual assault being seven times more likely to meet PTSD criteria compared to all other reported trauma types. These results, taken together, suggest that sexual assault is more likely to produce PTSD than other simple assaults.

Prevalence of other Psychological Reactions

Reactions to sexual assault other than PTSD are also common. Anxiety, anger, depression, substance abuse, and dissociation occur with high frequency and are discussed in the following sections.

Anxiety Anxiety and fear are among the most common reactions to sexual assault, but are often discussed under the rubric of PTSD. After an assault, women are often quite afraid of rape- and assault-related situations and also experience general anxiety at high levels. Such fears have been documented to last up to 16 years postassault. Indeed, in one study, only 23% of sexual assault victims did not show elevated fear 1 year after the assault. Although victims' fearfulness declined over time, they remained more fearful than nonvictimized controls 1 year after the assault. Thus, fear and anxiety levels remain elevated in many victims of sexual violence long after the assault.

Anger

Elevated anger levels are often seen among sexually assaulted women compared to nonvictimized controls. In a prospective study of 116 rape and other crime victims and 50 matched nonvictimized controls, Riggs et al. found that level of anger was related to aspects of the assault such as the use of a weapon and the victim's response to the attack. In addition, elevated anger levels were related to PTSD symptoms 1 month after the assault. Similarly, Feeny et al. found anger levels 1 month after an assault to be predictive of PTSD symptom severity at 3 months postassault. The authors hypothesized that anger interferes with recovery by preventing emotional processing of the traumatic event.

Depression Depression is another common reaction to sexual assault, although some evidence suggests that it is less persistent than fear and anxiety. In a study of psychological reactions to rape, Frank and Stewart found that almost half of their sample was diagnosed with major depression; however, over 3 months' time, these symptoms declined considerably. Retrospective examinations of depression among trauma victims, however, suggest more enduring symptoms. For example, one study found that an average of almost 22 years postassault, rape victims were more depressed than nonvictims.

In the North Carolina catchment area study, the lifetime prevalence rate of depression was significantly related to sexual assault: women who had been sexually assaulted were more than twice as likely to manifest a major depressive episode as women who had not been sexually assaulted. In addition, comorbidity of PTSD and depression is high: 48% of those who met lifetime criteria for PTSD also met lifetime criteria for depression, and 43% evidenced current comorbidity of the two disorders. In a sample seeking treatment for PTSD, 63% met criteria for current major depression. These results suggest that depression should be routinely assessed in victims of trauma who are seeking treatment.

Substance abuse Female assault victims, particularly those who develop PTSD, are also at risk for substance abuse. In a national probability sample of women, crime victims (sexual and nonsexual assault) were at increased risk of developing an alcohol use disorder, independent of their PTSD status. Moreover, those diagnosed with PTSD were 3 times more likely to have serious alcohol problems than those without the disorder and 14 times more likely than those without a history of criminal victimization. In the North Carolina catchment area study, lifetime prevalence of substance (alcohol and drugs) abuse and dependence was significantly related to the presence of sexual assault. In a recent longitudinal study, recent and past assault were associated with two to three times the risk of alcohol abuse in women, even after controlling for past alcohol use, age, race, and education.

Additionally, the prevalence of PTSD in patients enrolled in a substance abuse treatment program ranges from 12 to 34%, and prevalence rates are significantly higher among women, ranging from 33 to 59%. Among psychiatric outpatients, the rate of victimization among alcoholic women was higher

than among nonalcoholic women, suggesting a specific link between early victimization and later alcohol abuse among adult women. In summary, being a victim of a trauma appears to substantially increase the risk of developing a substance use disorder and vice versa: substance abuse seems to increase the likelihood of victimization.

Dissociation The construct of dissociation typically focuses on three clinical entities: alterations in memory (e.g., aspects of the trauma not consciously accessible), in identity (e.g., disengagement between self and environment), and in consciousness. Dissociation has been conceptualized to reflect emotional disengagement from trauma memories, which has been thought to hinder recovery. Indeed, several studies suggest that dissociation does hinder the processing of the traumatic event and subsequent natural recovery process. Trauma victims with PTSD exhibit more dissociative symptoms than those without PTSD. Moreover, if dissociation is present during or immediately after a trauma it is predictive of later psychopathology. Similarly, acute stress disorder, which emphasizes dissociative symptoms, has been found to be predictive of later posttrauma symptoms. The role of dissociation in promoting emotional disengagement needs further exploration in studies that evaluate risk factors in trauma exposed individuals.

Treatment of Sexual Assault-Related PTSD

Several cognitive behavioral treatment (CBT) programs have been developed to treat PTSD related to sexual assault, including exposure therapy, stress inoculation training (SIT), cognitive therapy (CT), eye movement desensitization and reprocessing (EMDR), and cognitive processing therapy (CPT). In this section, outcome studies, which evaluate the efficacy of these treatments with assault victims, are reviewed.

Recent studies examining the use of CBT with female sexual and physical assault victims have utilized well-controlled designs, and all but one utilized many of the methodological gold standards proposed by Foa and Meadows, including presence of PTSD diagnosis, assessment of PTSD severity, reliable and valid measures of related psychopathology, randomization to treatment conditions, and presence of control or comparison groups.

Prolonged exposure (PE) is an exposure therapy program that involves confrontation aimed at ameliorating PTSD, which typically includes mentally reliving the traumatic event repeatedly and *in vivo* confrontation with trauma-related situations, which evoke fear but are not objectively dangerous. In the first controlled study of exposure for rape-related PTSD, the efficacy of PE, SIT, supportive counseling (SC), and a wait-list control was compared for female victims of sexual assault. PTSD symptoms were assessed at pretreatment, posttreatment, and follow-up evaluations with psychometrically sound interviews and self-report measures administered by trained clinicians who were blind to treatment assignment. PE and SIT showed significant pre-post reductions on re-experiencing and avoidance clusters of PTSD, while SC and wait list did not. Also, at the end of treatment 50% of patients in SIT and 40% of those in PE no longer met criteria for PTSD; in contrast, only 10% of SC patients and none in the wait list no longer met diagnostic criteria. At follow-up, patients in the PE group tended to show further reduction in PTSD symptoms, whereas patients in SIT and SC did not show further improvement.

In a second study, Foa et al. obtained further support for the efficacy of a modified PE in treating rape-related PTSD. This study compared the efficacy of PE to SIT, a combination of both treatments (PE/SIT), and a wait-list condition. All active treatments resulted in substantial symptom reduction and were superior to the wait-list condition in ameliorating PTSD and related symptoms. PE was superior to the other treatments: it effected a greater reduction of anxiety and depression and resulted in fewer dropouts than SIT and PE/SIT. In addition, the number of patients who achieved good end-state functioning tended to be larger for the PE group than for the SIT and PE/SIT groups. In a third study, Foa et al. examined the efficacy of PE relative to a program that includes PE and cognitive restructuring (PE/CR). Results indicated that PE and PE/CR were superior to wait list at posttreatment on measures of PTSD and depression, with these gains maintained at follow-up. The addition of the cognitive restructuring component did not appear to enhance treatment outcomes; as with most posttreatment outcomes, PE and PE/CR did not differ from one another. Additionally, no differences in outcome were found among patients who were treated by experienced or recently trained counselors, suggesting that PE could be easily disseminated into community settings.

EMDR is a form of exposure accompanied by saccadic eye movements that has been the focus of considerable controversy due to initial claims of remarkable success in only a single session. Of the number of studies evaluating the efficacy of EMDR, only a few utilized well-controlled designs and thus yielded interpretable results. One well-controlled

study evaluated the efficacy of EMDR relative to a wait-list control condition for PTSD in female rape victims. Three sessions of EMDR resulted in greater improvement for PTSD symptoms (57% reduction in independently evaluated PTSD at posttreatment and 71% in self-reported PTSD at follow-up), relative to the wait-list condition (10% reduction at posttreatment). However, in contrast to other controlled studies, one therapist conducted all treatments, and therefore the contribution of therapist and treatment effects were confounded. A more recent study comparing PE and EMDR found that both treatment groups had significant reductions in PTSD symptoms at posttreatment. At 6-month follow-up, PE resulted in more adaptive scores on a number of anxiety and depression measures than those seen for participants who received EMDR. This study provides further support for the use of exposure techniques in the treatment of sexual assault.

Another recent study compared EMDR, PE, and relaxation training. This study found that PE effected the largest reductions in symptoms of re-experiencing and avoidance compared to EMDR and relaxation training and also effected the quickest reduction in avoidance symptoms. This study reported that EMDR and relaxation training did not differ from each other in efficacy or speed of change. Additionally, recent meta-analyses have found no evidence that the use of eye movements, integral to this treatment, have any effect on treatment outcomes. Thus, the use of the exposure techniques, which are a part of this treatment protocol, appears to account for the treatment outcomes using EMDR.

CPT is a program that was developed specifically for use with rape victims and includes cognitive restructuring and an exposure variant (i.e., writing the traumatic memories and reading the account aloud to the therapist). In a quasi-experimental design, CPT was compared to a naturally occurring wait-list control group. Overall, women who received CPT improved significantly from pre- to posttreatment (about 50% reduction in PTSD symptoms), whereas the wait-list group did not show significant improvement. In a second study, Resnick et al. compared the efficacy of a 12-session CPT, a 9-session PE, and a wait-list control in rape victims with chronic PTSD. After 9 weeks of PE and 12 of CPT, both treatments were highly effective in reducing PTSD symptoms. At follow-up assessments 3 and 9 months posttreatment, treatment gains were maintained.

A treatment specifically designed for women who experienced childhood sexual abuse has also been investigated; it is similar to CPT in that it is a combination treatment utilizing exposure and cognitive components. This treatment consists of skills training in affect and interpersonal regulation followed by a modified PE protocol (STAIR-modified PE). In a controlled study, relative to a wait-list condition, participants receiving STAIR-modified PE showed significant improvement in symptoms of PTSD, as well as improvements in affect regulation. This study provided evidence for the use of CBT procedures with childhood abuse survivors; however, future studies should compare the usefulness of the addition of the skills training as compared to PE alone, as previous studies utilizing PE have not found beneficial effects of adding components beyond exposure techniques.

In summary, the CBT program that has received the most empirical support in controlled studies is PE, and its efficacy in reducing PTSD symptoms is at least as high as that of comparison CBT programs. Additionally, recent evidence suggests that it can be successfully used in community settings. Thus, at present, PE is the recommended psychosocial treatment choice.

Predictors of PTSD

As is apparent from the results reviewed previously, not all women who have been sexually assaulted develop chronic PTSD; thus the identification of factors that are involved in hindering recovery has been the focus of much investigation. Foa and Riggs proposed to classify predictors into three categories: pretrauma, trauma, and posttrauma variables. In terms of pretrauma variables, it appears that poor psychological functioning prior to the trauma renders the individual vulnerable to developing chronic disturbances posttrauma. In addition, a prior history of traumatic events in child- or adulthood intensifies the response to subsequent trauma. Factors related to the trauma itself, including trauma severity, also appear to predict the development of PTSD, as well as posttrauma variables such as social support.

In a prospective study of 277 female assault victims, Amir et al. tested a predictive model of PTSD that included all three domains thought to be relevant to the prediction of PTSD. Several variables emerged as significant predictors at 1 and 3 months postassault: trauma history, trauma severity, cognitions and complex emotions, basic emotions, and early PTSD symptoms. Trauma history had an impact on cognitions and emotions, which, in turn, influenced PTSD symptom severity 4 and 12 weeks after the assault. Trauma severity influenced PTSD symptoms 2 weeks after the assault.

Recent meta-analyses provide additional information regarding the predictors of PTSD. One

meta-analysis found that the effects of predictors varied based on the sample under investigation; however, three variables were consistently significant predictors of PTSD: psychiatric history, childhood abuse, and family psychiatric history. Also, the largest effect sizes were found with predictors that were at or around the time of the trauma (e.g., social support following trauma). A second meta-analysis reviewed 476 studies that investigated PTSD predictors. Of the reviewed studies, 68 evaluated several of the hypothesized major predictors of PTSD, which include prior trauma, prior psychological adjustment, family history of psychopathology, perceived life threat, posttrauma social support, emotional response to trauma, and dissociation to trauma. Of these predictors, dissociation at the time of the trauma was the largest predictor of later developing PTSD. Indeed, the predictors with the largest effect sizes were those that were related more proximally to the trauma itself (e.g., dissociation and emotional response), similar to the finding in the Brewin et al. analyses. These results underscore the complexity of the mechanisms underlying the development of PTSD.

Conclusion

The majority of people in the United States are exposed to at least one traumatic experience in their lifetime; however, only a minority develop long-lasting psychological difficulties, including chronic PTSD. Importantly, not all traumas are equal in causing the development of PTSD; traumas inflicted by another person, especially sexual assaults, are particularly prone to produce lasting disturbances. Once established, PTSD is unlikely to remit unless treated. The most empirically validated type of therapy for this disorder is CBT, and among the various CBT programs, PE has gained the most empirical support.

See Also the Following Articles

Anger; Anxiety; Depression and Manic-Depressive Illness; Dissociation; Drug Use and Abuse; Trauma and Memory.

Further Reading

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental health disorders* (4th edn.). Washington, D.C.: American Psychiatric Association.

Davidson, J. and Foa, E. B. (eds.) (1993). *Posttraumatic stress disorder: DSM-IV and beyond*. Washington D.C.: American Psychiatric Press.

Foa, E. B. and Rothbaum, B. O. (1998). *Treating the trauma of rape*. New York: Guilford Press.

Foa, E. B., Dancu, C. V., Hembree, E. A., Jaycox, L. H., Meadows, E. A. and Street, G. P. (1999). A comparison of exposure therapy, stress inoculation training, and their combination for reducing posttraumatic stress disorder in female assault victims. *Journal of Consulting and Clinical Psychology* **67**(2), 194–200.

Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M. and Nelson, C. B. (1995). Posttraumatic stress disorder in the national comorbidity survey. *Archives of General Psychiatry* **52**, 1048–1060.

Resnick, H. S., Kilpatrick, D. G., Dansky, B. S., Saunders, B. E. and Best, C. L. (1993). Prevalence of civilian trauma and posttraumatic stress disorder in a representative national sample of women. *Journal of Consulting and Clinical Psychology* **61**(6), 984–991.

Sexual Behavior *See:* Sexual Dysfunction.

Sexual Dysfunction

W T O'Donohue and L Woodward Tolle
University of Nevada, Reno, Nevada, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by W T O'Donohue, L G Regev and D A Henderson, volume 3, pp 440–444, © 2000, Elsevier Inc.

Etiology
Interventions
Summary

Glossary

Anxiety	A symptom of stress.
Spectatoring	Monitoring one's sexual arousal, looking for signs of failure.
Stress	A subjective, phenomenological experience that is thought of by the individual experiencing it to be aversive; stress is generally thought to be a response to some external event or situation.
Stressor	External condition or situation that can cause a stress response in a given individual.

The *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.; DSM-IV-TR) characterizes sexual dysfunctions as either disturbances that occur in one of the four phases of the sexual response cycle or by pain that is associated with sexual intercourse. The four phases of the sexual response cycle are desire, excitement, orgasm, and resolution. The desire phase consists of fantasies about and the motivation to engage in sexual activity. The excitement stage includes a subjective sense of sexual pleasure and corresponding physiological changes, with major changes in males being tumescence and erection, and in females vasocongestion in the pelvis, vaginal lubrication and expansion, and the swelling of the external genitalia. The orgasm phase consists of a peaking of sexual pleasure accompanied by the release of sexual tension and contraction of the perineal muscles and reproductive organs. In males, ejaculation of semen occurs in this phase, and in females, there are rhythmic contractions of the wall of the vagina. Finally, the resolution phase involves a sense of muscular relaxation and well-being. Males are physiologically refractory to another erection and orgasm for a period of time; however females can respond to additional stimulation almost immediately. Sexual dysfunction can occur in any one of these phases or more than one, and factors such as age, experience, frequency and chronicity of the symptoms, subjective distress, and the symptoms' effects on other areas of functioning must be considered. The sexual dysfunction disorders as outlined in the DSM-IV-TR include sexual desire disorders (hypoactive sexual desire disorder and sexual aversion disorder), sexual arousal disorders (female sexual arousal disorder and male erectile disorder), orgasmic disorders (female orgasmic disorder, male orgasmic disorder, and premature ejaculation), sexual pain disorders (dyspareunia and vaginismus), sexual dysfunction due to a general medical condition, substance-induced sexual dysfunction, and sexual dysfunction not otherwise specified.

The prevalence of sexual dysfunctions is difficult to accurately report because, given the sensitivity of the issue, it is estimated that most cases go untreated or unreported. Of the epidemiological studies that have been conducted, there has been a large variability in terms of the prevalence of various sexual dysfunctions found in different populations. Other factors that might be contributing to these different findings could be the use of different assessment methods, the definitions used, or simply the varying characteristics of the samples used. Studies have estimated that the lifetime incidence of sexual problems is anywhere between 10 and 52% in males and 25 and 63% in females, whereas other studies have found ranges of only 5–13% in males and females.

Etiology

Research has demonstrated that the onset of sexual dysfunction varies greatly in that it could be present from the initial onset of sexual functioning or it could occur after a period of normal sexual functioning. Sometimes sexual dysfunction is not limited to certain types of stimulation, situations, or partners, and sometimes it is. Several considerations must be taken into account when making clinical judgments about the presence of a sexual dysfunction. Among these considerations are the individual's ethnic, cultural, religious, and social backgrounds. Another factor that adds a considerable degree of subjectivity to the diagnosis is the adequacy of sexual stimulation. There are no norms for this or for any of the factors clinicians must render judgments about. These factors may all influence individuals' sexual desire, expectations around sexual activity, and attitudes about their and their partner's performance.

There are also several risk factors associated with sexual dysfunctions that might be biological or psychological in nature. Biological factors can either directly lead to sexual problems or indirectly contribute. Some of the most common direct biological risk factors for sexual dysfunction include vascular disease, diabetes, hormone levels, the use of alcohol, and the use of certain medications (for example antihypertensives). Common indirect biological risk factors include age, cigarette smoking, chronic pain, heart disease, cancer, and other medical procedures that might alter an individual's appearance, leading to body-image insecurities and possible subsequent problems with normal sexual functioning. Psychological risk factors for sexual dysfunction include the presence of mood disorders, anxiety disorders, substance use disorders, eating disorders, and, largely, stress. Various negative life stressors adversely affect sexual functioning, including job stress, financial stress, bereavement or the loss of a loved one, divorce or relationship problems, problems conceiving, and marriage. Much of the literature examining the risk factors associated with sexual dysfunction in humans is correlational; however, studies using animals have provided more substantial evidence about the effects of stress on sexual functioning. Specifically, researchers have found that both physical stress (e.g., shock or immobilization) and social stress (e.g., overcrowding, other aggressive animals attacking, or isolation) result in copulatory disorder (equivalent to male erectile disorder in humans) in male mice and that, when faced with more mild stressors, mice copulatory behavior decreased. These findings are consistent in research conducted with human sexual behavior when faced with life stressors. Sexual dysfunction has been found to be more prominent among couples who report more severe and greater stressors in their lives compared to couples who report having fewer life stressors. In addition, sexual desire has been found to be lower or absent in individuals who report higher levels of occupational stress. Finally, research has discovered that sexual dysfunction is more likely in couples in which one or both partners have a diagnosed anxiety disorder.

Anxiety and stress, although appearing to frequently have different operational definitions, are frequently used interchangeably in the literature. Some studies examining the effects of anxiety on sexual arousal have produced the interesting findings that anxiety was found to negatively affect sexual arousal in individuals with sexual dysfunction but may actually enhance arousal in individuals with normal sexual functioning. It is possible that when sexually functional men and women become anxious during sexual activity, their autonomic arousal increases and they focus more on erotic cues, which enhances their arousal. Individuals with sexual dysfunction, on the other hand, become anxious, and when their autonomic arousal increases, they focus on not being able to perform (which Masters and Johnson called "performance anxiety" or spectatoring) and subsequently fail to become aroused.

When a sexual dysfunction is present, it often acts as an additional source of stress in the individual's or couple's lives, serving only to further exacerbate the dysfunction. This appears to be especially true in couples who are trying to conceive because there is not only pressure to perform sexually for pleasure and intimacy but also to get pregnant. Arguments, blame, embarrassment, and a lack of communication about sexual dysfunction can also increase stress and cause the sexual problem to remain unimproved.

Sexual Desire Disorders

The sexual desire disorders include hypoactive sexual desire disorder and sexual aversion disorder. Hypoactive sexual desire disorder is characterized by a deficiency or absence of sexual fantasies and desire for sexual activity, taking into consideration age and the individual's personal history. Sexual aversion disorder is characterized by a persistent or recurrent extreme aversion to and avoidance of all, or almost all, genital sexual contact with a sexual partner. Possible stress-related causes of both of the described disorders include anxiety over difficulty getting aroused or climaxing, marital stress, fear of intimacy, fear of losing control, guilt or religious prohibitions, fear of sexually transmitted diseases, fear of blood or pain during intercourse, posttraumatic stress disorder following rape or sexual assault, or a secondary trigger for panic disorder. Women are twice as likely to experience the desire disorders and are also more likely to report higher levels of psychological distress than men, more marital dissatisfaction than men, and greater self-reported stress levels than men, specifically in the areas of related to home and finances. Most frequently, hypoactive sexual desire disorder develops in adulthood, after a period of normal sexual functioning, in association with stressful live events or interpersonal difficulties.

Sexual Arousal Disorders

The sexual arousal disorders consist of female sexual arousal disorder and male erectile disorder. Female sexual arousal disorder is characterized by a persistent or recurrent inability to attain, or maintain until completion of the sexual activity, an adequate lubrication and swelling response of sexual excitement. Male erectile disorder is characterized by a persistent or recurrent inability to attain, or to maintain until

the completion of the sexual activity, an adequate erection. Women who experience female sexual arousal disorder often report little or no subjective sense of arousal and are unlikely to seek professional treatment because this problem can frequently be remedied with the aid of a lubricant. Without intervention, female sexual arousal disorder can result in painful intercourse, avoidance of sexual activity, and relationship problems. With male erectile disorder, different patterns can be seen with this sexual dysfunction. Some individuals may have difficulty obtaining an erection, some are able to obtain an erection but lose it when attempting penetration, some are able to sustain an erection on penetration but cannot on thrusting, and other males are able to sustain an erection only during self-masturbation or on awakening. Erectile difficulties are frequently associated with sexual anxiety, fear of failure, anxiety over sexual performance, and a decreased subjective sense of sexual excitement and pleasure. Erectile dysfunction can cause marital and relationship problems and may be the cause of infertility. The introduction of Viagra has caused male erectile disorder to become well known and in some cases obsolete.

Orgasmic Disorders

The orgasmic disorders consist of female orgasmic disorder, male orgasmic disorder, and premature ejaculation. Female orgasmic disorder is characterized by a persistent or recurrent delay in, or absence of, orgasm following a phase of normal female arousal taking into consideration the woman's age, sexual experience, and adequacy of sexual stimulation that she receives. Male orgasmic disorder is characterized by a persistent or recurrent delay in, or absence of, orgasm following a phase of normal male sexual arousal taking into consideration the man's age and his adequacy in focus, intensity, and duration. Premature ejaculation is characterized by a persistent or recurrent ejaculation with minimal stimulation before, on, or after penetration and before the individual wishes it. Stress-related etiological factors for women with female orgasmic disorder can include marital dissatisfaction, high levels of distress, and interpersonal stress. In addition, female orgasmic disorder has been found to adversely affect women's body image and self-esteem, which can then exacerbate the sexual dysfunction. Male orgasmic disorder is associated with performance anxiety, spectatoring, previous sexual traumas, and sexual and marital dissatisfaction. Premature ejaculation is associated with embarrassment and reluctance to engage in a sexual relationship due to fear of rejection and ridicule. Males with premature ejaculation often socially isolate themselves to prevent these events from taking place.

Sexual Pain Disorders

Sexual pain disorders include dyspareunia and vaginismus. Dyspareunia is characterized by recurrent or persistent genital pain associated with sexual intercourse in either a male or a female. Vaginismus is characterized by recurrent or persistent involuntary contraction of the perineal muscles surrounding the outer third of the vagina when penetration with a penis, finger, tampon or speculum into the vagina is attempted. Stress-related etiological factors associated with dyspareunia include anxiety, guilt from religious inhibitions, poor body image, and marital or relationship dissatisfaction. Vaginismus is associated with anxiety, interpersonal stress, and prior sexual abuse or traumatization.

Interventions

If stress is found to be an underlying cause of the sexual dysfunction, treatment will include strategies designed to decrease stress.

Psychoeducation

Psychoeducation can be a powerful tool in alleviating sexual dysfunction, not only in its ability to inform individuals but also to address false beliefs, unrealistic expectations, and lack of knowledge about physiology, anatomy, and normal sexual functioning, and it can, in and of itself, relieve anxiety associated with doing it wrong or feeling abnormal. As a result, psychoeducation can lead to rapid improvements in only a couple of sessions. Excerpts from human sexuality textbooks, complete with diagrams and pictures, can help provide a comprehensive overview of sexual functioning. Addressing common myths associated with human sexual functioning can also be helpful in educating individuals with sexual dysfunction.

Systematic Desensitization

Systematic desensitization is an effective therapy strategy designed to reduce anxiety. The individual is first equipped with relaxation skills and then is asked to come up with a hierarchy of increasingly anxiety-provoking stimuli or situations. When using this technique, an individual is placed in a deeply relaxed state and is presented with a series of gradually increasingly anxiety-provoking situations using imaginal exposure. The therapist starts with the least anxiety-provoking stimulus on the hierarchy and then slowly works up to more anxiety-inducing stimuli. The therapist asks the individual to imagine the stimulus, and when the individual experiences anxiety during the exposure exercise, the therapist asks the individual to allow him- or herself to get rid of the

image and then induces the original relaxed state. This procedure is used for longer and longer periods of time until the individual is able to retain the image without experiencing any anxiety; this is continued until the hierarchy is completed successfully.

Cognitive Restructuring

Intrusive negative thoughts can largely contribute to sexual dysfunctions. Ultimately, these thoughts interfere with sexual arousal and enjoyment because individuals are no longer able to concentrate on the sexual activity. Frequently, these negative thoughts may revolve around anxiety over their performance, one partner's dissatisfaction, or negative expectations about the sexual activity itself. In addition, individuals with multiple life stressors may experience thoughts concerning these stressors that may be a part of their sexual dysfunction and that can similarly disrupt sexual performance. In using cognitive restructuring, the therapist helps the individual identify these thoughts and then come up with other, more sensual or erotic thoughts to replace with these during intercourse. In this way, the individual can learn to turn the previously anxiety-provoking situation into a pleasurable one. In some cases, individuals have a difficult time coming up with the negative thoughts they typically have during intercourse. In these situations, it is useful to have them keep a diary, monitoring these thoughts as they occur and recording them following sexual activity. Another way to challenge individuals' negative thoughts is to help them determine whether the thoughts have any factual basis. This helps individuals let go of some negative and false assumptions that they might be carrying that are preventing them from enjoying intercourse.

Sensate Focus

Sensate focus is an effective tool used to help individuals with sexual dysfunction, specifically in cases in which performance anxiety appears to play a key role. Simply focusing on maintaining an erection or achieving an orgasm as the main goal of sexual intercourse can often cause anxiety to increase, particularly if there is a past history of the individual being unable to do so. Feelings of failure and disappointment can lead to increased anxiety about the next sexual encounter, and more preoccupation with achieving success often leads to subsequent failures. Sensate focus seeks to reduce anxiety by allowing individuals to take this pressure to perform out of the way by simply not allowing sexual intercourse for a time. A series of exercises are completed by the couple that include nonsexual touching and caressing while clothed, progressively moving through a series of increasingly demanding sexual activities over the course of treatment. The individuals are taught to focus on the sensations they experience rather than engaging in spectatoring. Because the pressure to aspire to the ultimate goal of maintaining an erection or achieving an orgasm is no longer the focus, the couple is able to restructure its goals in sexual encounters and thus reduce the anxiety and stress associated with performance anxiety.

Couples Therapy

Interpersonal, marital, and couples issues appear to play a central role in the development and/or maintenance of several sexual dysfunctions. Many of these issues are significantly improved by the implementation of communication skills training. Most sexual partners have a difficult time discussing their sexual relationship with one another. There are often fears about hurting one another's feelings, saying the wrong thing, or not being able to feel that they can ask for more of what they want and less of what they do not want. An improvement in overall communication can increase relationship satisfaction while creating a more effective line of communication to discuss more delicate matters, such as the couple's sexual relationship. Conflict resolution skills are important in maintaining a fair and even balance in times when there might otherwise be power struggles, and these can also help to build trust and respect. Enhanced communication not only alleviates stress from the relationship but also can assist in bringing the couple closer, both of which can contribute to a more satisfying sexual relationship.

Pharmacological Interventions

As mentioned previously, Viagra has become a household name, and it and its competitor medications are making a huge impact on sexual dysfunction. Not only does it significantly alleviate erectile dysfunction in a large proportion of men (the female equivalent to Viagra is beginning to be developed and has seen promising results in alleviating female arousal disorder), but it is also helping to decrease the social stigma surrounding sexual dysfunction and is allowing individuals to feel more comfortable about disclosing this problem to their physicians. The negative side to this is that, because it has become so convenient to get a prescription for this type of sexual enhancement drug, it is beginning to be abused by individuals who do not have sexual dysfunction and are simply using it recreationally. Naturally, physicians do not want to make it difficult for individuals who honestly may be suffering from erectile dysfunction to receive available support and help, so they do not withhold prescriptions from individuals reporting that they need the

drug. Thus, for now, no effective solution has been proposed to prevent this abuse from taking place.

Summary

Sexual dysfunctions are very common, affecting a large percentage of the population. Although epidemiological data are inconclusive, it is possible that only a small proportion of individuals suffering from sexual dysfunctions are being effectively treated. Stress appears to play a large role in the development, persistence, exacerbation, and maintenance of sexual dysfunctions. Several evidence-based interventions have been found to be effective with this population, including systematic desensitization, couples therapy, and sensate focus. In addition, pharmacological interventions have been helpful in alleviating the symptoms of sexual dysfunction. Further research is needed to continue to examine the role of stress in sexual dysfunctions and to development effective treatments to help prevent them.

Further Reading

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental health disorders* (4th edn.). Washington, D.C.: American Psychiatric Association.

Barlow, D. (ed.) (2001). *Clinical handbook of psychological disorders: a step by step treatment manual,* (3rd edn.). New York: Guilford Press.

Fasolo, C., Mirone, V., Gentile, V., et al. (2005). Premature ejaculation: prevalence and associated conditions in a sample of 12,558 men attending the andrology prevention week 2001 – a study of the Italian Society of Andrology. *Journal of Sexual Medicine* **2**(3), 376–382.

Heiman, J. R. (2002). Sexual dysfunction: overview of prevalence, etiological factors and treatments. *Journal of Sex Research* **39**(1), 73–78.

Lagana, L., Classen, C., Caldwell, R., et al. (2005). Sexual difficulties of patients with gynecological cancer. *Professional Psychology: Research and Practice* **36**(4), 391–399.

Leiblum, S., Brown, C., Wan, J., et al. (2005). Persistent sexual arousal syndrome: a descriptive study. *Journal of Sexual Medicine* **2**(3), 331–337.

Mohan, R. and Bhugra, D. (2005). Literature update: a critical review. *Sexual and Relationship Therapy* **20**(1), 115–122.

O'Donohue, W., Fisher, J. and Hayes, S. (eds.) (2003). *Cognitive behavior therapy: applying empirically supported techniques in your practice*. New York: John Wiley.

O'Donohue, W. and Geer, J. H. (1993). *Handbook of sexual dysfunctions: assessment and treatment*. Needham, MA: Allyn & Bacon.

Sexual Offenders

F M Saleh
University of Massachusetts Medical School, Worcester, MA USA
H M Malin
Institute for Advanced Study of Human Sexuality, San Francisco, CA, USA

© 2007 Elsevier Inc. All rights reserved.

General Considerations
Definitional Considerations
Treatment of Sex Offenders
Concluding Remarks

Glossary

Ephebophilia — A paraphilia in which the object of sexual attraction is a pubescent youth (child or adolescent), but not yet a fully sexually mature person, of either sex; also called hebephilia.

Paraphilia — Recurrent, intense sexual urges and fantasies, sometimes accompanied by sexually driven behaviors, occurring over a period of at least 6 months, involving one or more of the following: (1) a nonconsenting partner (e.g., children or a rape victim), (2) an object or body part (e.g., fur or feet), or (3) a sensory/emotional state (e.g., pain or humiliation) not typically associated with the archetypal sexual-response cycle. Paraphilias may be obligatory, in which case they are necessary components for completing the sexual-response cycle, or nonobligatory, in which case they may sometimes be incorporated in the sexual-response cycle but are not always essential for its completion. Paraphilias may be either pathological on nonpathological depending on the degree to which they impair the paraphile's social functioning and emotional well-being. They may be dangerous (such as lust

	murder or autoasphyxiophilia) or benign (such as transvestic fetishism or foot partialism) depending on their potential for lethality or adverse physical or social impact on self or others.
Pedophilia	A paraphilia involving sexual arousal to a prepubescent child or children of either sex.
Perversion	Once a psychiatric term roughly equivalent in meaning to paraphilia; now generally a pejorative term used by nonprofessionals to describe "kinky" sex or other expressions of sexuality (e.g., homosexual behavior) of which they disapprove.
Phenomenology	In the study of paraphilias or other psychiatric phenomena, the scientific method in which the phenomena experienced by the individual, as measured by observation and/or self-report, are described and classified rather than explained or interpreted.
Sex offender	A legal term denoting an individual who has broken a law governing the conduct of sexual behavior; the term is usually reserved for individuals who have been convicted of a crime involving sexual behavior.
Sexual-response cycle	The sequence of stages and events typically occurring in the progression of sexual activity from the beginning of sexual arousal to orgasm and resolution. Based on work by Masters and Johnson and amplified by other sexologists, the sexual-response cycle is sometimes arbitrarily divided into five stages or phases: (1) desire phase (sometimes called libido or drive); (2) excitement phase (sometimes called arousal), in which sexual arousal normally builds; (3) plateau (the point of highest arousal sustainable short of triggering orgasm); (4) orgasm (or climax); and (5) resolution (sometimes called the refractory period), in which arousal returns to prebaseline desire and during which an additional stimulus does not immediately lead to a repetition of the cycle.

General Considerations

It is universally assumed to be axiomatic that victims of sexual assault in its various manifestations suffer from the stress of their experiences to one degree or another. However, little or no scientific investigation has been conducted that sheds light on the relationship of stress to the experience of having been sexually victimized. Among important unstudied and, therefore, unanswerable questions are: What experiences are most likely to cause toxic stress, who is at greatest risk for developing stress-induced sequelae, and what factors might protect against or mitigate the adverse consequences of stress resulting from sexual victimization? It may be concluded, however, that those who work with and care for the sexually victimized are highly sensitized to the importance of stress-related components of sexual assault, whether the assault takes the form of uninvited genital exposure or rape with the threat of death.

The situation with respect to sex offenders is roughly equivalent, although there is the added variable that cultures typically do not concern themselves much with the impact of stress on sex offenders, who are generally regarded as willful evildoers and not as psychiatrically ill. In this article, it is not our intent to explore to any great degree the ongoing controversy of whether sex offenders are criminals or patients and the degree to which the psychiatric impairment in some sex offenders may influence culpability for their offending behavior.

The question of whether individuals who engage in illegal paraphilia-driven behaviors should be considered as having diminished criminal culpability given their psychiatric diagnosis is controversial and complex. Over a century ago, the eminent British psychiatrist Henry Maudsley expressed concern that society often fails to even consider such a possibility. He wrote, "If the law cannot adjust the measure of punishment to the actual degree of responsibility . . . that is no reason why we should shut our eyes to the facts; it is still our duty to place them on the record, in the confident assurance that a time will come when men will be more able to deal more wisely with them."

By definition, however, sex offenders are criminals with offenses ranging from misdemeanors to felonies. The term sex offender is strictly a legal term. There is no psychiatric diagnosis of sex offender, although there is some overlap between psychiatry and the law in that some individuals who have diagnosable psychiatric illness do commit sex offenses. Some sex offenders, including but not limited to those with paraphilias, may be more likely than others to commit sex offenses as a result of their psychiatric syndromes. Although severe comorbid psychiatric conditions may mitigate social response (e.g., the punishment may be less severe for an exhibitionist responding to command hallucinations), sex offenders are generally assumed to have adequate volitional control over their offending behaviors and are, thus, expected to conform their behavior to the expectations of society. Those who do not are likely to feel the full measure of retribution social systems are capable of bringing to

bear in dealing with them. That response, inevitably, induces toxic stress.

Indeed, with the diminishing emphasis on rehabilitation and the ascendancy of retribution in American penology, punishment is meted out with the specific intention of subjecting sex offenders to pain, suffering, and the attendant stress the retributive penal philosophy guarantees. The response of society to sex offenses is deliberately and systematically to induce intensive stress in the offender by prescribing a continuum of ever harsher punishment ranging, at minimum, from humiliation and social opprobrium to often-lengthy prison sentences and, occasionally, death. These stressors are intentionally inflicted with full knowledge of the adverse consequences of the stress they engender.

The underlying philosophy of retributive penology appears to be threefold. First, it is axiomatic that in serious offenses, retribution by the state is both appropriate and desirable because the state itself and the social values it defends are directly aggrieved by the crime. Second, it is axiomatic that the state has the duty to punish on behalf of individuals more specifically aggrieved by the crime, to bring miscreants to justice as their proxy, which prevents the degeneration of retribution into vengeance, which, in turn, risks sowing the seeds of vendetta and the further destabilization of society. Finally, retributive penology is motivated by the premise that institutionalized punishment promotes deterrence by example. That is, not only is the offender punished but those who would follow in his criminal footsteps, witnessing his punishment, will be deterred from committing similar crimes. In some cases, there is a bonus – a sex offender incarcerated for life or executed for his crimes will not himself recidivate. We mention, parenthetically, an observation first made by John Money that certain paraphilias are the only mental illnesses deemed to merit the death penalty rather than psychiatric treatment in some jurisdictions.

With respect to perpetrators of sex offenses, the authors are not aware of any peer-reviewed research or literature directly addressing the issue of stress in the offender, although it seems equally axiomatic that both victims and offenders may experience toxic stress when the sex offending takes place. There have been limited studies of stress in prison populations, which may provide leads for research in this area. It is thought, for example, that inmates or corrections officers with a certain psychological makeup or privileged position are immunized against the stress of the prison setting to a greater degree than other inmates or corrections officers in the same setting.

Part of the difficulty in assessing the impact of stress on both victims and sex offenders is that neither is a homogeneous group. In this article, we raise some issues relevant to stress in sex offenders and leave to others the parallel task of defining stress-related issues in the victims of sex offenses.

Definitional Considerations

Sex Offender

As we have pointed out, the term sex offense is a legal term. What constitutes a sex offense differs from jurisdiction to jurisdiction. As we pointed out in the preceding section, sex offenders are not a homogeneous group, with the sole exception of the fact that they have broken some law governing sexual behavior in a given culture or jurisdiction. It is sometimes not well understood that cross-cultural as well as intracultural considerations must be taken into account in defining sex offenses as well stress with respect to atypical sexual urges, fantasies, and behaviors.

For example, in some jurisdictions in the United States and in similarly less sexually tolerant foreign cultures, certain sexual behaviors such as premarital or extra-marital sexual contact, homosexual acts, cross-dressing, and heterosexual oral-genital or anal sexual contact are sex crimes and their practitioners are thus sex offenders. In virtually all cultures and jurisdictions, some sexual behaviors such as sexual contact with children (although the age of majority varies widely) are crimes and the individuals who engage in them are sex offenders as well.

Whether or not such criminal behaviors always, or even usually, engender stress depends on myriad considerations, including the degree of enforcement of the laws prohibiting the behavior, potential punishments, and whether individuals who engage in them find their own behavior personally acceptable (ego-syntonic) or unacceptable (ego-dystonic). Just as sex offenders are not a homogeneous group, the resultant stress from sex offending is not homogeneous.

Perhaps the only constant in the term sex offense and its equivalents in other cultures is that it is a legal construct. As such, it is not necessarily congruent with psychiatric or psychological views of these acts. Indeed, the differences between legal and psychiatric conceptualizations of sexual behaviors often leads to considerable dissonance and clouds the social debate about what should be done about sex offenders; they can lead to violent culture clashes, complicating considerably our understanding of stress as it relates to sex offending.

Phenomenology

Phenomenology, as we use the term here, is the personal, internal experience of the sexual fantasy or

behavior without respect to explanations for the behavior itself. Phenomenology, which has its roots in existential philosophy and, by extension, existential psychology, concerns itself with a scientific approach based on the phenomena themselves (here sexual drives, feelings, and motivated behaviors), without attempting to further explain them, as, for example, a psychodynamic approach might do. It also fails to embrace *in toto* the tenets and conventions of behaviorism. This theoretical approach has profound implications for treatment of "sex offenders."

Phenomenology is premised on the concept that behavior is determined by the way a person perceives reality rather than on some external objective reality. Thus, phenomenology does not rely on attempts to explain or interpret in order to classify sex offending. Nor does it rely overly much on changing behaviors themselves with the expectation that changing behaviors will change behavior. As a theory of psychology, phenomenology draws heavily on the tenets of existential psychology, which posits that psychology should concern itself with the organism's own experience of existence in the world. It draws heavily on the philosophical work of Karl Jaspers and the psychological application of that work by existential theorists such as Medart Boss, Ludwig Binswanger, and Viktor Frankl.

Paraphilias

Sex offenses and sex offenders are often classified as paraphilic or nonparaphilic.

The term, paraphilia, (from Greek *para-* beside + *philos* love) is an English translation of a term first proposed by Wilhelm Stekel in his book *Sexual Aberrations,* which first appeared in English about 1925. Reportedly, Stekel believed that a new term with less pejorative connotations than perversion (from Latin for turning around) would be helpful in examining and ameliorating these mental illnesses. Freud used the term, but it was not in widespread use in the psychiatric literature until the 1950s.

The *Diagnostic and Statistical Manual* (4th edn., text revision, DSM-IV-TR) of the American Psychiatric Association proposes that the "essential features of a paraphilia are recurrent, sexually arousing fantasies, sexual urges or behaviors generally involving 1) non-human objects, 2) the suffering or humiliation of oneself or one's partner, or 3) children or other non-consenting persons, that occur over a period of at least 6 months" (Criterion A; 2000: 566). DSM-IV-TR suggests that for some individuals "paraphilic fantasies or stimuli are obligatory for erotic arousal and are always included in sexual activity." For other individuals, paraphilic "preferences appear only episodically (e.g., during periods of stress)" (2000: 566). Criterion B for paraphilias, according to the DSM-IV-TR, is that they "must cause clinically significant impairment in social, occupational or other important area of functioning" (2000: 566). Thus, the DSM-IV-TR definition of paraphilia is at variance with the definition that we proposed previously, in that behavior alone is sufficient to make a diagnosis. We insist on a definition of paraphilia that presupposes motivated behavior that is sexually driven by recurrent, intense sexual urges and fantasies.

As an example, consider the variety of motivations driving the behavior of genital exposure:

- Preparing to bathe, dress, urinate, or otherwise engage in the gamut of biologically or socially based needs ordinarily considered to be nonsexual.
- Engaging in nude sunbathing, swimming, or other nonsexual recreational pursuits alone or with consenting others in lawful locations.
- Permitting a genital examination or other procedure for medical reasons.
- Engaging in sexual behavior with oneself or with consenting others.
- Attempting to attract the attention of nonconsenting people with or without the intention to initiate sexual contact.
- Attempting to shock, embarrass, humiliate, punish, frighten, or harass a nonconsenting person.
- Engaging in a social or cultural protest or making a political statement.
- Responding to a dare, exposing oneself as a penalty in a game or as a requirement for consensual initiation into social group.
- Engaging in a marginally socially accepted behavior, such as streaking or flashing, or participating exposed in a parade, festival, or other activity in which others are likely to do the same thing.
- Modeling for sculpture, painting, photography, or other artistic endeavors such as making films or participating in theatrical performances, either for remuneration or for altruistic reasons.
- Stripping for remuneration at a club or prenuptial party.
- Being sexually assaulted or intimidated into exposing oneself.
- Submitting to a lawful strip search.
- Preparing to commit a sexual assault on a nonconsenting person.
- Exposing oneself secondary to disinhibition caused by exogenous substances.
- Exposing oneself in response to command hallucinations.
- Exposing oneself secondary to dementia or traumatic brain insult.

Surely, most of these instances of genital exposure cannot be considered paraphilic. Some are prosocial, whereas some are antisocial. Some have negative or neutral erotosexual valence, whereas some are highly erotic. Paraphilic behavior, influencing as it does the sexual-response cycle, is by definition erotosexual. Yet only a subset of the exposing behaviors listed here that have positive erotosexual valence might be motivated by recurrent, intense sexual urges and fantasies outside some presumptive or defined norm. Only those that satisfy this criterion can be considered paraphilic.

Etiology There is no known etiology of paraphilias, but some investigators believe there are associations between paraphilias and other life experiences. The most common association reported is between paraphilias and early childhood experiences. Theories change and track dominant schools of psychological thought. For example, when psychodynamic theory was the predominant psychiatric theory, fur fetishism was explained as a sequella of the birth experience, in which an infant became fixated on the pubic hair of the mother during childhood. In the light of present-day understanding, such an explanation seems unlikely. However, more satisfying explanations have not yet been forthcoming.

More recently, various researchers have postulated a relationship between the development of paraphilias and childhood physical or sexual abuse. Such theories have led to concepts such as the cycle of abuse, in which it is postulated that the abused becomes the abuser by some mechanism as yet not clearly delineated. There is currently considerable debate about the proposition that such abuse may have occurred in preverbal children and the memories recovered many years later. Empirical studies have not been kind to the cycle of abuse construct, although it is at the foundation of a number of still widely used treatment regimens.

Although a good statistical correlation between having been abused as a child and perpetrating abuse as an adult has not been demonstrated, in some sexually abusive individuals there is a considerable history of childhood sexual, physical, and/or emotional abuse or neglect. The recalled or sometimes corroborated sexual biographies of a great many paraphilic individuals contain experiences in which the patient's paraphilia seems closely linked to these experiences. For example, individuals who received frequent spankings, enemas, or even highly embarrassing physical examinations as children have sometimes reported that as adults these unpleasant experiences have become highly pleasurable with positive erotic valence and may be incorporated as obligatory or nonobligatory paraphilic components of their sexual-response cycle. One possible explanation for this transformation may be the mechanism of opponent process theory of acquired motivation proposed by Richard Solomon and his colleagues, in which, like skydiving, an intensely fear-laden experience gradually becomes highly pleasurable, sought-after exhilarating activity. Other investigators have suggested that paraphilias are reinforced by masturbation to deviant fantasies. Unfortunately, all these ideas, as well as others, await solid empirical evidence. At this point, we can say with certainty only that the genesis of paraphilias remains largely unknown.

Prevalence There are in excess of 50 named paraphilias in the sexological literature. Some appear to be quite rare, whereas others, especially in their nonobligatory form, seem relatively common. The paraphilias most frequently coming to the attention of clinicians are pedophilia (which typically includes ephebophilia, although they are different paraphilias), voyeurism, and exhibitionism, according to DSM-IV-TR. This is hardly surprising, given that acting on these paraphilias is illegal in most jurisdictions, and individuals acting on these paraphilias are the most likely to seek clinical treatment, typically mandated by the judiciary. Given the distribution of pornographic materials with legal paraphilic content, for example, films and magazines devoted to transvestic fetishism, leather fetishism, or the milder forms of sadism and masochism, we might speculate that these paraphilias are among the more common. Yet sadism and masochism are relatively infrequently seen in clinical practice.

Erotic materials depicting other kinds of paraphilias, for example zoophilia, are typically suppressed. And although zoophilia is illegal in most jurisdictions, it seems to be less aggressively prosecuted, even when it comes to the attention of the authorities, than, say, sexual behavior involving children. It is difficult to extrapolate from any available data the prevalence of a given paraphilia or an all-inclusive prevalence of paraphilias in the American culture. DSM-IV-TR reports that, except for sexual masochism, with a sex ratio of 20 males to each female, paraphilias are almost never diagnosed in females. DSM-IV-TR further reports that approximately half of all individuals with paraphilia seen in clinics are married.

Treatment of Sex Offenders

In practice, the treatment of sex offenders may depend less on careful differential diagnosis than on the convictions that the treatment provider holds about sex offenders. Treatment providers with a behavioral view rely more heavily on psychological treatments

designed to change behaviors than do treatment providers who subscribe to a more medical model approach. These providers tend to use more pharmacologically based interventions than their behaviorist colleagues. The authors rely on an approach that takes advantage of an appropriate mix of interventions, following a careful differential diagnosis.

Psychological Treatments

The various psychological therapies, whether individual or group therapies, are designed to change behaviors through a variety of strategies including challenging patients' distorted thought and belief patterns (cognitive behavioral therapy), developing strategies to anticipate and resist deviant sexual cravings (relapse prevention), and replacing maladaptive behaviors with more functional behaviors (behavioral therapy). More traditional psychodynamic and psychoanalytic psychotherapy treatment modalities have been largely abandoned in the treatment of sex offenders because of their lack of demonstrated efficacy despite decades of attempts.

Behavior therapy Behavioral therapy is guided by learning theory, specifically social learning theory. In contrast to psychodynamic therapies, behavior therapies are primarily concerned with the aberrant behavior itself and not its underlying cause. That is, behavior therapy aims to decrease and extinguish deviant sexual arousal through a set of techniques, such as systematic desensitization, aversion therapy, biofeedback, minimal arousal conditioning with aversive conditioning, and covert sensitization. (Masturbatory satiation, a form of arousal reconditioning, is no longer widely used.)

Covert sensitization, a form of aversion therapy, has been widely used in the treatment of paraphilic patients. It is postulated that the pairing of deviant sexual fantasies with the mental image of its humiliating consequences (i.e., the aversive experience) will dissuade a patient from acting out sexually. In covert sensitization, deviant arousal or excitement is paired with a visualized distressing thought or fantasy. To give an example, a voyeur will be asked to imagine being taking into custody as soon as he begins to contemplate peeping into other people's windows.

In contrast to the aversive behavioral therapies, operant conditioning methods use positive reinforcers. For example, positive olfactory conditioning employs pleasant aromas with nondeviant sexual stimuli.

Cognitive-behavioral therapy Cognitive-behavioral therapy hypothesizes that paraphilias are maintained by distorted cognitions and reinforced by masturbation to inappropriate fantasy. The aim of cognitive therapy is to change maladaptive thoughts and beliefs through techniques such as cognitive restructuring and thought stopping and to eliminate reinforcers that maintain distorted cognitions. The technique of cognitive restructuring teaches patients to identify, challenge, and ultimately replace erroneous (i.e., distorted) beliefs with adaptive and pro-social cognitions. Thought stopping is a technique that aims to decrease the frequency and duration of deviant sexual thoughts.

More often than not, a combination of one or more of these therapies is used. Victim empathy training has traditionally been an important component of all behavioral treatment and involves helping a patient take on the perspective of his victim(s). In addition, assertiveness and a variety of social skills training programs have been deemed to be important components in treatment programs, although recent research seems to indicate that empathy or social skills training have little or no impact on sexual recidivism in treated populations.

Relapse prevention A relapse is a recurrence of undesirable behavior after a period that has been free of the undesirable behavior. The term "relapse" is not synonymous with "recidivism" which indicates re-arrest for an illegal behavior. The objective of relapse prevention is to help a patient maintain behavioral changes. In order to reduce the risk of inappropriate sexual behavior, patients are taught to identify their deviant cycle and to develop strategies to implement when the cycle is triggered.

Group therapy Group psychotherapy is a widely accepted treatment modality used in the treatment of paraphilic and nonparaphilic sex offenders. It incorporates relapse prevention and cognitive-behavioral techniques, as well as empathy and social skills training. Group therapy creates a supportive milieu that is conducive to frank discussion and that provides guidance regarding effective relapse prevention strategies. It also encourages the development of a social support network and of healthy adult relationships. One of the primary advantages of group therapy over individual therapy is that it allows patients to be challenged by members of the group rather than by the therapist. This is important given that patients may respond and be more willing to change if challenged by one of their peers.

Effectiveness of psychological treatments The effectiveness of the psychological treatments in treating paraphilias is difficult to assess. One measure of treatment success has traditionally been recidivism. Some studies have concluded that these treatments

have positive effects; other studies have found minimal or no effect with behavioral therapies. The interpretation of these studies may be confounded by a number of variables such as non-standardized treatment delivery or heterogeneous populations. Also, studies of the treatment of sex offenders generally do not differentiate paraphilic from nonparaphilic offenders.

Nor is it clear that treatment delivered in one setting, say a prison, generalizes to another setting, such as the community. State-dependent learning, for example, may impact psychological treatments that depend on a psycho-educational approach. Relapse prevention, a widely used psychological treatment modality, may be less effective than previously believed. In one randomized clinical trial that compared the reoffense rates of patients in an inpatient relapse prevention treatment program with the reoffense rates in two untreated prison control groups, no significant differences among the three groups in rates of sexual or violent reoffending over an 8-year follow-up period were found.

Biological Treatments

The authors believe that with certain patients, for example those with dangerous paraphilias or those at acute risk of victimizing someone, pharmacological intervention should be the treatment option of first resort. Biologically based treatments should also be made available to paraphilic patients if, for example, cravings for deviant sexual activities become overpowering or when specific symptoms are not responsive to less invasive treatment modalities (e.g., behavioral therapy).

Surgery Among the biologically based treatments, we must differentiate between surgical and pharmacological interventions. Stereotaxic hypothalamotomy is only of historical interest. On the other hand, the orchiectomy (surgical removal of the gonads) data are particularly relevant in that they provide the basis for our understanding of the benefits of the currently used testosterone-lowering agents. Orchiectomy consistently lowers the rate of recidivism. Orchiectomy has been rendered largely unnecessary with the advent of the pharmacological treatment.

Pharmacotherapy For ethical and medical reasons, randomized, double-blind, placebo-controlled pharmacological treatment studies cannot be conducted with symptomatic paraphilic patients who reside in the community. Such studies could only be justified on patients with relatively innocuous paraphilias. Nevertheless, the number of medications used to treat paraphilias has been steadily increasing. These medications can be divided into two general classes: those that act by lowering available androgens (e.g., progesterone derivatives and the gonadotropin releasing hormone analogs) and those that do not directly reduce androgen activity but reduce libido and/or sexual function as a side effect (e.g., some serotonin reuptake inhibitors). As with all pharmacological treatments, the choice of which medication to use is primarily based on the patient's presenting symptoms, his or her concomitant psychiatric/neurological disorders, and the results of the psychosexual and medical workup. A regrettable shortcoming of pharmacotherapy, however, is that it does not selectively reduce the risk for sex offending without generally impairing overall sexual functioning.

Concluding Remarks

There are a number of studies in the literature suggesting that treatment can be helpful, particularly in paraphilic samples. One such study reported rates of recidivism following treatment in a primarily paraphilic sample that were well below the rates reported in the world literature for comparable untreated cases. In another, a meta-analysis that reviewed data from 43 studies examining the sex offense recidivism rate for mixed sex offenders without differentiating between paraphilic and nonparaphilic patients found that recidivism, on average, was approximately 12% lower for treatment groups than for nontreatment groups.

Measuring treatment effectiveness presents significant problems. Recidivism, defined as some subset of involvement with the criminal justice system during or after the completion of treatment, is the most commonly used end point for determining treatment effectiveness in sex offenders. But not all paraphiles are sex offenders and not all sex offenders are paraphiles. Clearly, it is not possible to use recidivism as a measure of treatment efficacy with patients whose paraphilias are not illegal. Defining recidivism and dealing with important but largely unknown variables such as the base rate for reoffending and other methodological impediments greatly influence the interpretation of the available data and are central to answering questions about which treatment modalities work and which do not.

Sex offenses are prevalent and those motivated by paraphilia-driven phenomenology often constitute significant psychiatric disorders. Other, nonparaphilic sex crimes may have their roots in serious psychopathology as well. All too frequently, patients manifesting these conditions are simply ostracized, stigmatized, and punished when what is called for is enlightened psychiatric care.

Ignoring phenomenology and defining even observable manifestations of human sexuality in purely behavioral terms can only lead to egregious errors. Nowhere is this more apparent than in the arena of sex offender treatment. The phrase itself seems to us to be misleading, if not an oxymoron. With or without civil commitment statutes, clinicians are often called on to treat sex offenders. For example, a phenenomenological view of rape implies that rape is not a monolithic sex offense but a highly variable one. Rape may be an opportunistic act, carried out by an individual without a sense of conscience. It may be the result of command hallucinations. It may also be paraphilic. It seems reasonable to assume, as in all psychiatric treatment, that how we approach the treatment of a rapist, or for that matter any sex offender, must depend to some degree on differential diagnosis, which, in turn, depends on the phenomenology underlying the behavior.

See Also the Following Articles

Antisocial Disorders; Sexual Assault.

Further Reading

Abel, G. G., Mittleman, M., Becker, J. V., et al. (1988). Predicting childmolesters' response to treatment. *Annals of the New York Academy of Sciences* **528**, 223–224.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th edn., text rev.). Washington, DC: American Psychiatric Association.

Berlin, F. S. (1989). The paraphilias and depo-provera: some medical, ethical and legal considerations. *Bulletin of the American Academy of Psychiatry and the Law* **17**(3), 233–239.

Berlin, F. and Malin, H. (1991). Media distortion of the public's perception of recidivism and psychiatric rehabilitation. *American Journal of Psychiatry* **148**(11), 1572–1576.

Bradford, J. M. W. (2000). The treatment of sexual deviation using a pharmacological approach. *Journal of Sex Research* **37**(3), 248–257.

Freund, K. (1980). Therapeutic sex drive reduction. *Acta Psychiatrica Scandinavica* **62**(supplement 287), 1–39.

Frost, J. J. and Mayberg, H. S. (1986). Alteration in brain opiate receptor binding in man following arousal using C-11 carfentanil and positive emission tomography. *Journal of Nuclear Medicine* **27**, 1027 (abstract of presentation at the 33rd annual meeting of the Society of Nuclear Medicine).

Greenberg, D. M., Bradford, J. M. W., Curry, S., et al. (1996). A comparison of treatment of paraphilias with three serotonin reuptake inhibitors: a retrospective study. *Bulletin of the American Academy of Psychiatry and the Law* **24**(4), 525–532.

Hanson, R. K., Gordon, A., Harris, A. J., et al. (2002). First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sex Abuse* **14**(2), 169–194.

Krafft-Ebing, R. von. (1965). *Psychopathia sexualis (1886)*. New York: G. P. Putnam's Sons.

Langelüddecke, A. (1963). *Die Entmannung von Sittlichkeitsverbrechern in Deutchland*. Berlin: DeGruyter.

Laws, D. R., Hudson, S. M. and Ward, T. (2000). *Remaking relapse prevention with sex offenders: a sourcebook*. Thousand Oaks, CA: Sage.

Marshall, D. R. and Laws, W. L. (2003). A brief history of behavioral and cognitive behavioral approaches to sexual offenders. Part 1: Early developments. *Sexual Abuse: A Journal of Research and Treatment* **15**(2), 93–120.

Money, J. (1985). *The destroying angel*. Buffalo, NY: Prometheus Books.

Money, J. (1986). *Venuses penuses: sexology, sexosophy and exigency theory*. Buffalo, NY: Prometheus Books.

Moser, C. (2002). Are any of the paraphilias in the DSM mental disorders? *Sexual Behavior* **31**(6), 490–491.

Murphy, W. D. and Carich, M. S. (2001). Cognitive distortions and restructuring in sexual abuser treatment. In: Carich, M. S. & Stephens, E. M. (eds.). *Handbook for sexual abuser assessment and treatment*, pp. 65–75. Brandon, VT: Safer Society Press.

Ortman, J. (1980). The treatment of sexual offenders: castration and antihormone therapy. *International Journal of Law and Psychiatry* **3**, 443–451.

Quinsey, V. L., Harris, G. T., Rice, M. E., et al. (1993). Assessing treatment efficacy in outcome studies of sex offenders. *Journal of Interpersonal Violence* **8**, 512–523.

Rosler, A. and Witztum, E. (1998). Treatment of men with paraphilia with a long acting analogue of gonadotropin-releasing hormone. *New England Journal of Medicine* **338**, 416–465.

Saleh, F. M. and Berlin, F. S. (2003). Sex hormones, neurotransmitters and psychopharmacological treatments in men with paraphilic disorders. *Journal Child Sexual Abuse* **12**(3–4), 53–74.

Saleh, F. M. and Guidry, L. L. (2003). Psychosocial and biological treatment considerations for the paraphilic and nonparaphilic sex offender. *Journal of the American Academy of Psychiatry and the Law* **31**, 486–493.

Saleh, F. M., Niel, T. and Fishman, M. (2004). Treatment of paraphilia in young adults with leuprolide acetate: a preliminary case report series. *Journal of Forensic Sciences* **49**(6), 2–6.

Stürup, G. K. (1968). Treatment of sexual offenders in Herstedvester, Denmark: the rapist. *Acta Psychiatrica Scandinavica* **44**, 5–63.

Shift Work *See:* Sleep Loss, Jet Lag, and Shift Work.

Sickle Cell Disease and Stress

K Midence
Institute of Medical and Social Care Research, Bangor, UK

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 445–446, © 2000, Elsevier Inc.

Medical Complications
Stress and Chronic Illness
Sickle Cell Disease-Related Stressors
Coping Styles and Sickle Cell Disease

Glossary

Sickle cell disease (SCD)	An inherited blood disorder including sickle cell anemia, sickle cell hemoglobin disease, and sickle β-thalassaemia. SCD affects around 1 in 400 African Americans, and approximately 50,000 individuals have sickle cell anemia in the United States. SCD is also found in Europe, India, Africa, Central and South America, and the Caribbean.
Vaso-occlusive episodes	Recurrent and painful crises are caused by the sickle hemoglobin molecules, which may stick together under low oxygen tension, distorting the red blood cells into a pathonogmonic and rigid sickle shape and reducing their normal elasticity.

SCD, like all chronic illnesses, can be considered as a major stressor to which children and their families have to adjust by reducing the stresses associated with the condition and developing appropriate coping skills to ensure that the affected children and their families can lead a normal life.

Medical Complications

Vaso-occlusive crises caused by sickled red cells are responsible for chronic and acute damage to tissue and organs and cause recurrent and unpredictable episodes of acute pain, often very violent, in the affected parts of the body, most usually the arms, legs, joints, and back, causing 91% of hospital admissions. Other complications include abdominal pain and enlargement of the liver or spleen, infections, delayed growth and onset of puberty, problems with pregnancy, priapism, avascular necrosis, and strokes, which may also occur in childhood. Effective clinical management requires medical and psychological skills. Medical treatment is mainly supportive and involves rapid and adequate analgesia, hydration, and treatment of any precipitating events such as infection and inflammatory conditions.

Stress and Chronic Illness

Wallander and Varni suggested that families with a chronically ill child are at greater risk of maladjustment because they are exposed to a greater number of stressful situations. These may happen to the child directly (e.g., missing school because of complications) or indirectly, affecting the family (e.g., marital disagreements and financial difficulties). They also suggested a set of factors associated with adjustment to chronic illness, including intrapersonal (e.g., functional independence), interpersonal (e.g., coping style of the family), and social-ecological (e.g., family functioning and socioeconomic status). This approach emphasizes the need to identify the specific stresses related to a disease and also generates the hypothesis that increased frequency and severity of stressors will result in poor adjustment. Adjustment could then be defined in terms of how well an individual deals with the typical stressful situations associated with a specific chronic illness.

Sickle Cell Disease-Related Stressors

For children and adolescents with SCD, the most frequent and severe stress is probably the experience of pain, both mild and severe; the uncertainty of painful crises; its adverse effect on their daily lives; and the risk of a sudden death. Pain crises and hospitalizations have been reported by children with SCD as

the most stressful aspects of their lives. SCD can disrupt school attendance and activities (e.g., sports), which may have a negative effect on their social life and academic performance. In adults with SCD, adjustment has been related more closely to life stress (e.g., financial and work) and social support. SCD can limit or disrupt a wide range of activities because exercise, cold exposure, and other factors can precipitate sickling.

However, family life and employment are the most significant areas of social life in adults with SCD. An increased personal responsibility for coping and financial status play important roles in coping with SCD, and most of those affected belong to ethnic groups that are already at a significant socioeconomic disadvantage. It has been suggested that the more psychosocial stressors within the family, the less effectively the SCD patient and the family cope with the condition and that families with more social support available appear to cope more favorably with the stressors associated with SCD.

Psychological research suggests that life stress and psychosocial adversity are the most likely precursors of maladjustment and psychopathology in adults with SCD. Thompson and colleagues found that over half of adults with SCD in their study met the criteria for poor adjustment, with 40 and 32% showing symptoms of depression and anxiety, respectively. Better adjustment was associated with lower levels of perceived daily stress, higher efficacy expectations about illness tasks, and less use of palliative coping and negative or passive coping strategies, but it was not associated with type of SCD or measures of severity. Family functioning was related to psychological status, with higher family support and lower family conflict and control among those with better adjustment.

Coping Styles and Sickle Cell Disease

Research has shown that active coping strategies (i.e., addressing stressors) are associated with good adjustment, whereas avoidance coping strategies (i.e., ignoring or avoiding stressors) are associated with poor adjustment. Passive strategies and negative thinking have been associated with lower activity levels and greater treatment requirements. Thompson and colleagues provided a transactional stress and coping model that views chronic illness as a potential stressor to which the individual and the family attempt to adapt. Thus, the transactional stress and coping model includes parameters that are illness-specific as well as common demographic and adaptational parameters that are assessed across illnesses.

Coping is hypothesized to be a function of the transactions of illness paramenters, demographic parameters, and adaptational processes. Illness parameters include type, severity, and age of onset. Demographic parameters include patient age, gender, and socioeconomic status. Adaptational processes include cognitive processes of stress appraisal; expectations of efficacy; and health locus of control, self-esteem, and causal attributions.

Thompson and coworkers found that the pain-coping strategies of children with SCD were characterized by negative thinking, which accounted for a 20% increment. Variables of the transactional stress and coping model accounted for 30% of the variance in child-reported symptoms and 45–49% of the variance in mother-reported behavior problems. Thompson and colleagues also found that higher levels of stress and lower levels of self-worth accounted for a significant portion of the variance in self-reported symptoms.

Future research needs to investigate illness-specific patterns of adaptation and to develop interventions aimed at enhancing methods of managing cognitive appraisals of illness-related stress.

See Also the Following Articles

Avoidance; Coping Skills.

Further Reading

Midence, K. and Elander, J. (1994). *Sickle cell disease: a psychosocial approach*. Oxford: Radcliffe Medical Press.

Midence, K. and Elander, J. (1996). Adjustment and coping in adults with sickle cell disease: an assessment of research evidence. *British Journal of Health Psychology* **1**, 95–111.

Thompson, R. J., Gil, K. M., Abrams, M. R., et al. (1992). Stress, coping, and psychological adjustment of adults with SCD. *Journal of Consulting and Clinical Psychology* **60**, 433–440.

Thompson, R. J., Jr., Gil, K. M., Burbach, D. J., et al. (1993). Role of child and maternal processes in the psychological adjustment of children with sickle cell disease. *Journal of Consulting and Clinical Psychology* **61**, 468–474.

Thompson, R. J., Jr. and Gustafson, K. E. (1996). *Adaptation to chronic childhood illness*. Washington, DC.: American Psychological Association Press.

Thompson, R. J., Jr., Gustafson, K. E., Gil, K. M., et al. (1998). Illness specific patterns of psychological adjustment and cognitive adaptational processes in children with cystic fibrosis and sickle cell disease. *Journal of Clinical Psychology* **54**, 121–128.

Wallander, J. L. and Varni, J. W. (1992). Adjustment in children with chronic physical disorders: programmatic research on a disability-stress-coping model. In: La Greca, A. M., Siegel, L. J., Wallander, J. L. & Walker, C. E. (eds.) *Stress and coping in child health*. New York: Guilford Press.

Sleep Loss, Jet Lag, and Shift Work

E Van Cauter
University of Chicago, Chicago, IL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E V Cauter, volume 3, pp 447–448, © 2000, Elsevier Inc.

Stress and Sleep Loss
Stress and Jet Lag
Stress and Shift Work

Glossary

Endogenous circadian rhythm	An oscillation with a period of approximately 24 h generated by a pacemaker mechanism located in the suprachiasmatic nuclei of the hypothalamus. This endogenous rhythm persists in the absence of any periodic variations in the environment (e.g., light–dark cycle and rest–activity cycle).
Jet lag syndrome	The malaise experienced by the majority of travelers during the days following a flight across time zones. The syndrome usually includes symptoms of daytime fatigue, difficulties initiating and maintaining sleep, deterioration of mood, vigilance, and cognitive function and may also involve gastrointestinal disorders. The syndrome is believed to result from the misalignment between bodily rhythms and the 24-h periodicities in the environment.
Sympathetic and parasympathetic tone	The autonomous nervous system plays a crucial role in the stress response and includes the sympathetic and parasympathetic nervous systems. Nearly all organs receive both sympathetic and parasympathetic inputs. During stress, sympathetic activation is increased and thus the balance of sympathetic versus parasympathetic control (i.e., the sympatho-vagal balance) is shifted toward higher sympathetic as compared to parasympathetic control.

The evolution of industrial society has been associated with major changes in the behavioral control of the sleep–wake cycle and in the environmental light–dark cycle. The availability of electrical lighting has made extended activity hours and around-the-clock operations possible. Misalignments among endogenous circadian rhythms, sleep–wake cycle, and light–dark cycle are highly prevalent in the working population and also affect millions of travelers who use rapid air transportation across time zones. While it is the common perception that sleep loss, jet lag, and shift work are stressful, there are remarkably few studies of either subjective or objective markers of stress in these conditions.

Stress and Sleep Loss

Voluntary sleep curtailment is a hallmark of modern society. In the United States, normal sleep duration has decreased from approximately 8.5 h in 1960 to an average of less than 7.0 h in 2006. To maximize the time available for work and leisure activities, it has become socially desirable to curtail sleep to the shortest amount tolerable. Around-the-clock operations imply chronic partial sleep loss for the large population of shift workers and the repeated occurrence of total sleep deprivation for key personnel.

Although multiple studies have shown that total sleep deprivation or repeated sleep curtailment by as little as 2–3 h per night results in marked decrements in mood and performance, subjective indices of stress in subjects studied in the laboratory remain generally low. In particular, self-ratings on scales of tense and nervous remain stable across extended periods of total sleep deprivation under comfortable and sedentary laboratory conditions. Similarly, ratings of perceived stress remain low and stable in subjects submitted to up to 1 week of partial sleep deprivation in the laboratory. Sleep deprivation does not appear to be a classical stressor, resulting in an immediate activation of the sympathetic nervous system or in the rapid release of stress hormones. Indeed, the only known immediate effect of acute total sleep deprivation on physiological function is the absence of responses normally elicited by sleep onset or awakening from sleep, such as declines and increases in heart rate. However, sleep loss, even when only partial, appears to be associated with an alteration in the regulation of the hypothalamic-pituitary-adrenal (HPA) axis, the major neuroendocrine transducer of stress. Indeed, in sleep-deprived subjects, the normal morning-to-evening decline of blood levels of the stress hormone cortisol is slowed down and, as a result, evening cortisol levels are elevated. This alteration suggests that the rate of recovery of the HPA axis from exogenous stimulation by stressors may be affected by sleep loss. An elevation of cardiac

sympathetic activity occurs following recurrent sleep restriction. Thus, sleep loss may indeed result in an activation of the major biological mechanisms involved in the stress response. However, even though sleep loss appears to be interpreted biologically as a stressor, subjectively, the subjects do not seem to experience sleep restriction as stressful since they consistently rate their perceived stress as low, irrespective of bedtime duration. This apparent dissociation between the biological and psychological response to sleep loss may only occur in the unchallenging environment of the laboratory. The combination of increased sympathetic nervous activity and elevated evening cortisol levels in a state of sleep debt may play a role in the deterioration of parameters of glucose tolerance that is associated with short sleep.

Stress and Jet Lag

Jet lag syndrome is a condition of desynchrony between the near 24-h (i.e., circadian) endogenous rhythm generated in the suprachiasmatic nuclei of the hypothalamus and the external light–dark and rest–activity cycles that arises following rapid transportation across time zones. The syndrome usually includes symptoms of fatigue, sleep disturbances, reduced cognitive and psychomotor performance, and gastrointestinal tract disorders. The syndrome subsides as adaptation to local time is achieved. The 24-h secretory rhythms of hormones of the HPA axis adapt at a slower rate (averaging 1 day per hour of time difference) than diurnal variations in catecholamine levels or sympatho-vagal tone. Adaptation to jet lag nearly always involves some degree of sleep loss and therefore alterations of HPA regulation typical of sleep deprivation are to be expected.

Stress and Shift Work

To meet the demands of around-the-clock operations, millions of shift workers live in a condition of chronic sleep debt and desynchrony between endogenous rhythms and environmental periodicities. Indeed, studies have shown that shift workers get significantly less sleep per scheduled sleep period than daytime workers, with many sleeping on average less than 5 h per work day. In contrast to occasional jet lag, shift work is a major health hazard, involving an increased risk of cardiovascular illness, gastrointestinal disorders, psychosocial complaints, sleep disturbances, substance abuse, reduced immune function, and infertility.

Shift workers thus live in a situation of conflicting zeitgebers, which almost never allow a complete shift of the circadian system. Indeed, several studies have shown that workers on permanent or rotating night shifts do not adapt to these schedules, even after several years. In particular, the temporal profile of secretion of the stress hormone cortisol shows little, if any, adaptation to the work schedule. There is evidence for increased blood pressure and increased sympathetic versus parasympathetic tone during night shift as compared to day shift activity in emergency rooms of hospitals. Under chronic conditions, such disturbances could result in cardiovascular illness.

See Also the Following Articles

Industrialized Societies; Sleep, Sleep Disorders, and Stress; Workplace Stress.

Further Reading

Adams, S. L., Rose, D. M., Weiss, J., et al. (1998). Ambulatory blood pressure and monitoring of emergency physicians before, during, and after a night shift. *Academic Emergency Medicine* **5**, 871–877.

Bliwise, D. L. (1996). Historical change in the report of day-time fatigue. *Sleep* **19**, 462–464.

Bonnet, M. and Arand, D. (1995). We are chronically sleep deprived. *Sleep* **18**, 908–911.

Broman, J. E., Lundh, L. G. and Hetta, J. (1996). Insufficient sleep in the general population. *Neurophysiologie Clinique* **26**, 30–39.

Czeisler, C. A., Johnson, M. P., Duffy, J. F., et al. (1990). Exposure to bright light and darkness to treat physiologic maladaptation to night work. *New England Journal of Medicine* **322**, 1253–1259.

Dinges, D., Pack, F., Williams, K., et al. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 hours per night. *Sleep* **20**, 267–277.

Ekstrand, P., Bostrom, P. A., Arborelius, M., et al. (1996). Cardiovascular risk factors in commercial flight aircrew officers compared with those in the general population. *Angiology* **47**, 1089–1094.

Folkard, S., Minors, D. S. and Waterhouse, J. M. (1985). Chronobiology and shift work: current issues and trends. *Chronobiologia* **12**, 31–54.

Knutsson, A., Akerstedt, T., Orth-Gomer, K., et al. (1986). Increased risk of ischaemic heart disease in shift workers. *Lancet* **2**, 89–92.

Leproult, R., Copinschi, G., Buxton, O., et al. (1997). Sleep loss results in an elevation of cortisol levels the next evening. *Sleep* **20**, 865–870.

Roden, M., Koller, M., Pirich, K., et al. (1993). The circadian melatonin and cortisol secretion pattern in permanent night shift workers. *American Journal of Physiology* **265**, R261–R267.

Spiegel, K., Leproult, R., L'Hermite-Balériaux, M., et al. (2004). Impact of sleep duration on the 24-hour leptin

profile: relationships with sympatho-vagal balance, cortisol and TSH. *Journal of Clinical Endocrinology and Metabolism* **89**, 5762–5771.
Spiegel, K., Knutson, K., Leproult, R., et al. (2005). Sleep loss: a novel risk factor for insulin resistance and type 2 diabetes. *Journal of Applied Physiology* **99**, 2008–2019.
Tepas, D. I. and Carvalhais, A. B. (1990). Sleep patterns of shiftworkers. *Occupational Medicine* **5**, 199–208.
Weibel, L., Spiegel, K., Follenius, M., et al. (1996). Internal dissociation of the circadian markers of the cortisol rhythm in night workers. *American Journal of Physiology* **270**, E608–E613.

Sleep, Sleep Disorders, and Stress

A N Vgontzas, S Pejovic and M Karataraki
Penn State University College of Medicine, Hershey, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A N Vgontzas, E O Bixler and A K Kales, volume 3, pp 449–457, © 2000, Elsevier Inc.

Normal Sleep
Sleep Disorders
Sleep, Sleep Disorders, and Physical and Emotional Stress
Sleep and Stress System
Conclusion

Glossary

Addison's disease	An endocrine disease that is associated with aldosterone and cortisol (hormones) by the adrenal glands.
Apnea	Cessation of breathing lasting at least 10 s.
Cataplexy	A brief and sudden loss of muscle control without loss of consciousness usually precipitated by strong emotions.
Continuous positive airway pressure (CPAP)	A small apparatus which through a soft plastic mask blows air gently through the nose into the throat to keep the airway open.
Cosinor analysis	Least square approximation of time series using a cosine function of known period.
Dihydroxyphenylacetic acid (DOPAC)	A metabolite of dopamine.
Dihydroxyphenylglycol (DHPG)	A metabolite of norepinephrine produced by the action of the enzyme monoamine oxidase (MAO) which can be measured in urine.
Hypersomnia	Disorder of excessive diurnal and nocturnal sleep.
Hypnagogic	Occurrence of an event while drifting into sleep.
Hypoxia	A pathological condition in which the body as a whole (generalized hypoxia) or region of the body (tissue hypoxia) is deprived of adequate oxygen supply.
Idiopathic	Of unknown causation.
Microneurography	A method to record the traffic of impulses in human nerves with percutaneously inserted needle electrodes.
Narcolepsy	Sleep disorder associated with excessive daytime sleepiness, involuntary daytime sleep episodes, disturbed nocturnal sleep, and cataplexy (sudden loss of muscle tone).
Parasomnia	Is any sleep disorder such as sleepwalking, teeth grinding, night terrors, rhythmic movement disorder, rapid eye movement (REM) behavior disorder, restless leg syndrome, and sleep talking, characterized by partial arousals during sleep or during transitions between wakefulness and sleep.
Polysomnography	The continuous and simultaneous recording of physiological variables during sleep, e.g., electroencephalography, electrooculography, electromyography, electrocardiography, respiratory flow, etc.
Pulsatile analysis	A global analysis of the mechanical function of the cardiovascular system, including the determinants of cardiac output.
Slow-wave sleep (SWS)	Sleep characterized by electroencephalogram waves of duration slower than 4 Hz and high amplitude.
Stage rapid eye movement (REM) sleep	The stage of sleep with highest brain activity characterized by enhanced brain metabolism and vivid dreaming.

Sleep appears to be an important element in maintaining equilibrium or homeostasis that is vital for the survival of living organisms. In this article, we review the association between sleep, sleep disorders, and stress, which has been defined as a state of disharmony, or threatened homeostasis.

Normal Sleep

Humans spend at least one-third of their lives asleep, yet there is little understanding of why we need sleep and what mechanisms underlie its capacities for physical and mental restoration. However, there has been a significant increase of empirical knowledge that is useful in the evaluation and management of most sleep complaints and their underlying disorder.

The interaction of circadian effects, i.e., usual time to go to sleep and amount of prior wakefulness (homeostatic response), determines the onset and amount of sleep. Natural sleep–wake rhythms cycle at about 25 h rather than coinciding with the solar 24-h schedule. As a result, many people depend on external cues to keep their diurnal cycle on time. The normal diurnal clock resists natural changes in its pattern by more than about 1 h per day, which explains the sleep difficulties that usually accompany adaptation to new time zones or switches in work shifts.

Individuals differ considerably in their natural sleep patterns. Most adults in nontropical areas are comfortable with 6.5–8.0 h daily, taken in a single period. Children and adolescents sleep more than adults, and young adults sleep more than older ones. Normal sleep consists of four to six behaviorally and electroencephalographically (EEG) defined cycles, including periods during which the brain is active (associated with rapid eye movements, called REM sleep), preceded by four progressively deeper, quieter sleep stages graded 1–4 on the basis of increasingly slow EEG patterns. Slow-wave sleep (SWS; also referred to as deep sleep; stages 3 and 4) gradually lessens with age and usually disappears in the elderly.

Sleep Disorders

Sleep disorders are common in the general population and are associated with significant medical, psychological, and social disturbances. Insomnia, the most common sleep disorder, most often reflects psychological disturbances. Excessive daytime sleepiness is the predominant complaint of most patients evaluated in sleep disorders clinics and often reflects organic dysfunction. Narcolepsy, idiopathic hypersomnia, and sleep apnea are the most common disorders associated with excessive daytime sleepiness. Narcolepsy and idiopathic hypersomnia are chronic brain disorders with an onset at a young age; sleep apnea occurs predominantly in middle-aged men and (to a lesser degree) women and is associated with obesity and cardiovascular complications. The parasomnias, including sleepwalking, night terrors, and nightmares, have benign implications in childhood but often reflect psychopathology or significant stress in adolescents and adults and organicity in the elderly.

Sleep, Sleep Disorders, and Physical and Emotional Stress

"Stress is life and life is stress" in Hans Selye's words. If it is a fact that there is no life without stress, it is also true that there is no life without sleep. Thus, it is not surprising that the association between sleep and physical and emotional stress has been long recognized. Hippocrates observed early on the relationship between mental and physical health and sleep and stated that sleeplessness is a sign of pain and suffering and may lead to mental illness while sleeping during the day (daytime sleepiness) is an indication of illness. Also, sleep has been an ancient remedy in the hands of physicians to combat emotional or physical stress, e.g., infection.

Acute or short-term sleep disturbance is usually associated with a variety of situational stresses (work-related, interpersonal, or financial difficulties) or with medical problems such as pain, cardiopulmonary or gastrointestinal disorders, or the febrile prodromes to influenza. Various drugs, including those that humans have been using for hundreds of years to decrease stress and dysphoria such as alcohol, nicotine, and caffeine, can affect adversely the quantity and quality of sleep. Stressful life events at the onset of sleep disorders are quite common, particularly in those disorders that appear to be of emotional origin, such as insomnia. It has been found that 75% of the chronic insomniacs had experienced some stressful life event at about the time insomnia began. Stressful life events are present, although to a lesser degree compared to insomniacs, in other sleep disorders, such as narcolepsy or sleep apnea, or adult sleepwalking, night terrors, and nightmares. The presence of stressful life events at the onset of organic disorders, such as narcolepsy or sleep apnea, is considered coincidental. However, it is interesting to note that, for example, it has been proposed that in narcolepsy it is the combination of genetic factors and stress, including emotional stress, which leads to the manifestation of the disorder. In this respect, a recent study showed that a majority of people with narcolepsy had a stressful life event at onset of their illness.

Psychological distress appears to be quite common among patients with sleep disorders. In some of these disorders, such as chronic insomnia or adult parasomnias, it appears to be primary or etiologic. In other sleep disorders, such as sleep apnea, narcolepsy, or idiopathic hypersomnia, it appears to be secondary or reactive to the chronic physical problems associated with the sleep disorder. Furthermore, it has been reported that in insomniacs, it is not only stress, but also how they handle stress that leads to chronic difficulties in initiating and maintaining

sleep. In particular, insomniacs handle external stress and conflict by internalization of their emotions which results in a combination of emotional arousal and physiological activation. This mental and physiological hyperarousal leads to difficulty in initiating sleep, whether at the beginning of the sleep period or when returning to sleep following awakening. Fear of sleeplessness and performance anxiety further intensifies the emotional arousal. All of these factors through behavioral conditioning lead to a vicious circle that perpetuates insomnia.

Because of the high frequency of emotional distress associated with sleep disorders, stress management techniques are routinely recommended to all sleep-disorder patients. Specifically, in insomnia, besides recommending sleep hygiene measures, medication, or psychotherapy, a series of stress management measures are routinely recommended, such as: recognizing the association between stressful events and sleeplessness; ventilating conflict and anger to avoid internalization of emotions; addressing daily worries a few hours before bedtime; becoming tolerant of occasional sleeplessness; avoiding rumination over sleep difficulty; and relaxation techniques. Increasing empirical evidence suggests that cognitive behavioral therapy (CBT) for insomnia is the most effective non-pharmacological approach in the management of chronic insomnia. Besides using educational and behavioral techniques, CBT involves cognitive restructuring that aims at identifying and altering faulty beliefs and attitudes about sleep that in themselves are believed to perpetuate the vicious cycle of insomnia, fear of sleeplessness, emotional arousal, and further sleep disturbance. However, the CBT approach does not address long-term personality patterns such as internalization of psychological conflict and inhibition of emotions associated with insomnia. In the sleep disorders of organic origin, such as narcolepsy or sleep apnea, supportive therapy, including educating and advising patients and their families gently but honestly about the nature of the chronic disorder and that the symptoms are not under voluntary control, is of utmost importance. Furthermore, helping the patient to adjust his or her lifestyle to the changes imposed by the sleep disorder, i.e., sleep or cataplectic attacks in narcolepsy, or enhancing his/her compliance to the chronic use of cumbersome and often, anxiety-provoking therapies, such as continuous positive airway pressure (CPAP) in sleep apnea, is an important dimension of the overall management of the sleep-disorder patient.

Sleep disturbance is a very frequent symptom in most of the psychiatric disorders. Depression, anxiety disorders, and acute onset schizophrenia are frequently associated with sleep disturbance. Depression, in particular melancholia, is associated with early morning awakening, while sleep loss is the early sign of an acute onset schizophrenia or an early sign of relapse. In contrast, sleep disturbance or sleep loss can lead to mental illness. For example, it has been reported that chronic insomnia is associated with a higher incidence of major depression. Also, acute sleep loss can lead to a manic episode in a patient with bipolar disorder. Thus, it is not surprising that sleep therapy or hypnotherapy was very common in the practice of mental health professionals in the nineteenth and early twentieth centuries. Today, the improvement of the quality and quantity of sleep of a person under emotional distress is one of the therapeutic priorities of the practicing clinician.

Sleep and Stress System

In mammalian organisms, including humans, the stress system consists of components of the central nervous system (CNS), including: the corticotropin-releasing hormone (CRH) neurons of the hypothalamic paraventricular nucleus; and several, mostly norepinephric nuclei of the brain stem, and their peripheral limbs, the hypothalamic-pituitary-adrenal (HPA) axis and the peripheral autonomic system, whose main function is to maintain homeostasis, both in the resting and stress states.

Normal Sleep and Stress System

Although the association between stress and sleep has been noticed for hundreds of years, a more systematic and scientifically based approach in the relation between sleep and stress system has only existed since the 1980s. It was in 1983 that Dr. Weitzman and his colleagues reported that sleep, in particular SWS, appears to have an inhibiting influence on the HPA axis and cortisol secretion (**Figure 1**). Since then, several studies have replicated this finding. In a reverse mode, activation of the HPA axis and central administration of CRH can lead to arousal and sleeplessness. In normal sleepers, wakefulness and stage 1 sleep (light sleep) accompany cortisol increases, while SWS is associated with declining plasma cortisol levels. In addition, in healthy people, induced sleep disruption (continuous arousals) is associated with significant increases of plasma cortisol levels. Furthermore, mean plasma cortisol level was significantly higher in a group of subjects with a shorter total sleep time than another group with a longer total sleep time. Finally, it has been suggested that the effect of aging on the levels and diurnal variation of human adrenocorticotropic activity could be involved in the etiology of poor sleep in the elderly.

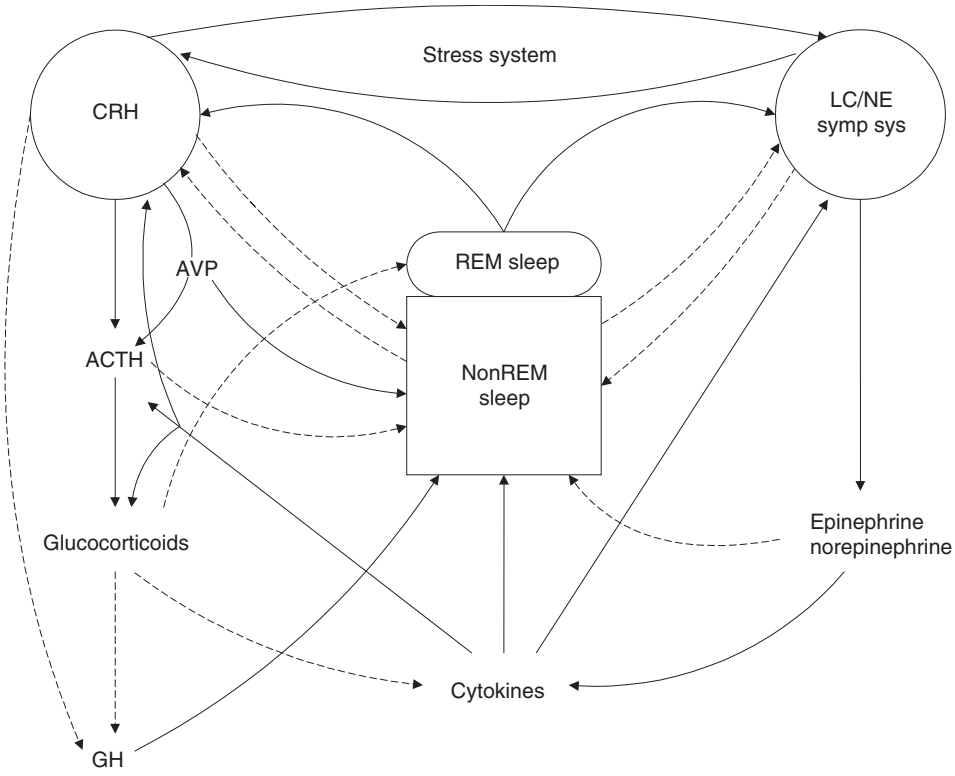


Figure 1 A simplified, heuristic model of the interactions between central and peripheral components of the stress system with sleep and REM sleep. CRH–ACTH cortisol and LC/NE systems are two main arousal systems that their activation leads to wakefulness. NonREM sleep is associated with an inhibitory effect on these two wakefulness promoting systems, whereas during REM there is an activation of both limbs of the stress system. ACTH, corticotropin; AVP, arginine vasopressin; CRH, corticotropin-releasing hormone; GH, growth hormone; LC/NE symp sys, locus ceruleus-norepinephrine/sympathetic system. A solid line denotes promotion/stimulation; a broken line denotes disturbance/inhibition.

REM sleep, which is a state of CNS activation that resembles unconscious wakefulness (paradoxical sleep), appears to be associated with a higher activity of the HPA axis (Figure 1). An early study showed that urinary 17-hydroxy corticoids were increased during REM epochs in urological patients. One study in healthy, normal sleepers showed that the amount of REM sleep was positively correlated with 24-h urinary free-cortisol (UFC) excretion. These results are consistent with the co-existence of stress system activation, and REM sleep increases in patients with melancholic depression. In rats, chronic mild stress resulted in disruption of REM sleep, including a reduced latency to the onset of first REM period, providing support for the validity of the proposed association between stress and mechanisms underlying endogenous depression, including REM sleep alterations.

Although the effects of sleep on the sympathetic nervous system have been studied less frequently, it appears that in humans during sleep, catecholaminergic activity is decreased (Figure 1). In particular, catecholaminergic activity appears to be suppressed during SWS, while there is an activation of the sympathetic nervous system during REM sleep. It has been found that in healthy, normal sleepers, the amount of REM sleep correlated to 24-h excretion of urinary catecholamines while sympathetic-nerve activity measured using microneurography, was elevated during REM sleep. Interestingly, in rats, locus ceruleus (LC) neurons fire most frequently during waking, less during nonREM (NREM) sleep, and are silent during paradoxical sleep.

The stress system is closely linked with immune responses. In particular, three pro-inflammatory cytokines (tumor necrosis factor-α (TNF-α), interleukin (IL)-1, and IK-6) have a strong stimulating effect on HPA axis (Figure 1). Glucocorticoids, the endproducts of the HPA axis, inhibit the production of all three cytokines, while catecholamines, the other products of the stress system, stimulate IL-6 secretion. In animals, several studies have demonstrated the potential role of these cytokines, particularly IL-1 and TNF-α in sleep regulation. In humans, these cytokines appear to be increased at nighttime. A study in young, healthy, normal sleepers showed that the amount and

quality (depth) of sleep correlated negatively with the overall daytime secretion of IL-6. These data suggest that a good night's sleep is associated with decreased secretion of IL-6, a good sense of wellbeing, and that good sleep is associated with decreased exposure of tissues to the pro-inflammatory and potentially detrimental actions of IL-6.

The role of corticotropin-releasing hormone in sleep–wake regulation CRH, which is produced and released from parvicellular neurons of the paraventricular nucleus, is the key regulator of the HPA axis. Release of CRH is followed by enhanced secretion of adrenocorticotropic hormone (ACTH) from the anterior pituitary and cortisol from the adrenal cortex. CRH exerts various influences on behavior, including wakefulness and sleep.

Several studies in animals have shown that intracerebroventricular administration (ICV) of CRH induces increased waking. Also, specific CRH receptor blockade reduces spontaneous awakening and rat strains deficient in the synthesis and secretion of hypothalamic CRH spend less time awake than do control counterparts. Finally, CRH concentration in the brains of rats exhibits circadian fluctuations and is highest when rats are most active. Several reports indicate that CRH is excitatory in the LC, hippocampus, cerebral cortex, and some portions of the hypothalamus. Spontaneous discharge rates in the LC are highest during arousal and lowest during sleep. Also, it has been reported that the brain β_1-adrenergic system is involved in the restraint stress-induced increase in arousal in rats.

In humans, the majority of studies suggest that the sleep of young individuals is rather resistant to the arousing effects of CRH. In contrast, middle-aged individuals responded to an equivalent dose of CRH with significantly more wakefulness and suppression of SWS compared to baseline. Based on these findings, we concluded that middle-aged men show increased vulnerability of sleep to stress hormones, possibly resulting in impairments in the quality of sleep during periods of stress. These findings suggest that changes in sleep physiology associated with middle-age play a significant role in the marked increase of prevalence of insomnia in middle-age. Also, peripheral administration of CRH is associated with REM suppression, which is stronger in the young than in the middle-aged. The administration of ACTH and its analogs in humans has been associated with general CNS activation consisting of a decreased sleep period time and sleep efficiency, and an increase of sleep latency. Continuous administration of ACTH produced a marked reduction in REM sleep.

Arginine vasopressin (AVP) is a peptide which acts as a co-factor of CRH in the activation of the HPA system. Several experiments in rats have suggested that the continuous administration of AVP causes sleep disturbance.

Steroid effects on sleep The administration of glucocorticoids has been reported to be associated with a robust suppression of REM sleep (Figure 1). Also, in some studies, it has been reported that the continuous or pulsatile nocturnal administration of cortisol paradoxically was associated with a modest increase of SWS in addition to the well-established decrease of REM sleep. The opposite effects of cortisol on SWS compared to CRH are believed to be mediated by a feedback inhibition of CRH from cortisol. In contrast, the glucocorticoid antagonist, mifepristone (RU-486), led to a significant worsening of sleep in normal controls. The administration of mineralocorticoids is not associated with significant changes of sleep structure.

A study in patients with Addison's disease recently demonstrated that an evening replacement dose of hydrocortisone was necessary for proper expression of REM sleep (*vide infra*) suggesting that glucocorticoids have some permissive action for this sleep parameter, possibly reflecting an inverse u-shaped dose–response curve.

Interestingly, in clinical practice, the use of glucocorticoids is associated with sleep disturbance. In fact, in a multicenter study in which steroids were used on a short-term basis, sleep disturbance was one of the most common side effects.

Aging, HPA axis, and sleep Old age is associated with marked sleep changes consisting of increased wakefulness, minimal amounts of SWS, declining amounts of REM sleep, and earlier retiring and rising times. Some studies have shown that older adults have elevated cortisol levels at the time of the circadian nadir and have higher basal cortisol levels than younger adults. It is difficult to discern whether the latter changes are associated with aging or increased medical morbidity common in this group. It has been suggested that the effect of aging on the levels and diurnal variation of human adrenocorticotropic activity could be involved in the etiology of poor sleep in the elderly. Higher evening cortisol concentrations are associated with lower amounts of REM sleep and increased wakefulness. More recently, it was shown that older women without estrogen replacement therapy (ERT), when subjected to mild stress, showed greater disturbances in sleep parameters than women receiving ERT, suggesting that female hormones protect women's sleep from external stressors.

Sleep Deprivation and Stress System

If sleep is important for our sense of wellbeing, then it is conceivable that sleep deprivation represents a stress to humans and should be associated with activation of the stress system. However, several studies that have assessed the effects of one night's sleep deprivation on the HPA axis have shown that cortisol secretion is either not or minimally affected by sleep following prolonged wakefulness. Two studies have reported somewhat antithetical results, with the one study showing that cortisol secretion is elevated the next evening following sleep deprivation, and the other study showing a significant decrease of plasma cortisol levels the next day and recovery night. The study that showed decrease of plasma cortisol levels following sleep loss indicated that this inhibition of the HPA axis activity was associated with an enhanced activity of the growth hormone axis. The finding that sleep deprivation leads to lower cortisol levels postdeprivation (primarily during the subsequent night of sleep) suggests that lowering the level of HPA activity, which is increased in depression, may be the mechanism through which sleep deprivation improves the mood of depressed individuals. Prolonged sleep deprivation that results in death is associated with increased plasma norepinephrine levels, and higher ACTH and corticosteroid levels at the later phase of sleep deprivation. It is postulated that these increases are due to the stress of dying rather than to sleep loss.

A recent study demonstrated that a week of partial sleep restriction did not affect significantly the 24-h cortisol secretion either in men or women. However, there was a significant effect in terms of the circadian secretory pattern of the hormone as indicated by the significant time effect, and the time and gender interaction effect. Specifically, the peak cortisol secretion in the morning was lower after sleep restriction than baseline and this difference was stronger in men than in women. In both genders, the peak of cortisol secretion was shifted by 2 h earlier than baseline (08.00 versus 06.00). Sleep restriction did not affect significantly the nadir value of cortisol secretion, neither the afternoon, nor evening presleep levels.

The inconsistency in cortisol secretion after total or partial sleep loss may be explained by methodological differences in these studies, i.e., nonstressful deprivation is not associated with increased cortisol secretion. Also, the higher evening cortisol levels reported in studies where sleep was curtailed at the beginning of the night may represent a change of the 24-h-secretory pattern, e.g., a shift of the nadir point later at night.

Although sleep deprivation appears to have beneficial effects on the mood of depressed individuals, in healthy individuals, sleep deprivation leads to next-day fatigue and somnolence, decreased concentration and attention, and increased vulnerability to accidents. These effects appear to be mediated partially by the increased levels of IL-6, which is related to the stress system. It has been shown that following sleep deprivation, there is an increase of IL-6 levels during the daytime and a decrease during the first night of sleep following sleep deprivation. The increased levels of IL-6 during the daytime can explain the somnolence and fatigue experienced by the individuals after one night's sleep loss, whereas the nighttime undersecretion of IL-6 might be responsible for the better quality (depth) of their sleep. The exact mechanisms of this elevation are not known. However, alterations of the stress system following sleep deprivation may be responsible for this elevation. For example, the decreased levels of cortisol postdeprivation and/or possible increase of catecholaminergic activity can lead to elevation of plasma IL-6 levels. It has been shown that sleep deprivation in humans appears to enhance norepinephrine and dopamine activity.

Sleep Disorders and Stress System

Although sleep disorders/disturbances with their various physical and mental effects on the individual should be expected to affect the stress system, information regarding the effects of sleep disorders/disturbances on this system is lacking.

Insomnia and stress system Insomnia, a symptom of various psychiatric or medical disorders, may also be the result of an environmental disturbance or a stressful situation. When insomnia is chronic and severe, it may itself become a stress that affects the patient's life so greatly that it is perceived by the patient as a distinct disorder itself. Either way, as a manifestation of stress or a stressor itself, insomnia is expected to be related to the stress system. There is a paucity of data regarding the activity of the stress system in insomniacs. Few studies have measured cortisol levels in poor sleepers or insomniacs, and those results are inconsistent. The majority of these studies reported no difference between controls and poor sleepers in 24-h cortisol or 17-hydroxy steroid excretion. In a study in 15 young adults with insomnia, it was demonstrated that 24-h UFC levels were positively correlated with total wake time (Table 1). In addition, 24-h urinary levels of catecholamines and their metabolites, dihydroxyphenyl glycol (DHPG), and 3,4-dihydroxy-phenylacetic acid (DOPAC) were positively correlated with percent stage 1 sleep and wake time after sleep onset. Norepinephrine tended to correlate positively with percent stage 1 sleep and

Table 1 Sleep disorders and the stress system

Insomnia	Increased 24-h secretion of ACTH-cortisol, more intense in the evening; positive correlations between 24-h UFC and urinary catecholamines and polysomnographic indices of severity of the disorder
Sleep apnea	Increased excretion of urinary catecholamines; increased nighttime plasma catecholamines; increased daytime plasma IL-6 and TNF-α levels; mild hypercortisolemia compared to obese controls
Narcolepsy	Increased daytime plasma IL-6 and TNF-α levels
Idiopathic hypersomnia	Mild hypocortisolism; enhanced ACTH secretion to exogenous CRH; central CRH deficiency

ACTH, adrenocorticotropic hormone; CRH, corticotropin-releasing hormone; IL-6, interleukin-6; TNF-α, tumor-necrosis factor-α; UFC, urinary free-cortisol.

wake time after sleep onset and negatively with percent SWS. It was concluded that in chronic insomnia, the activity of both limbs of the stress system, i.e., the HPA axis and the sympathetic system, relates positively to the degree of objective sleep disturbance. However, in this study also, the total amount of the 24-h UFC or catecholamines excretion was not different from normative values. The results on the positive correlation between urinary catecholamines and polysomnographic indices of sleep disturbance in chronic insomnia are consistent with previous studies that showed elevated urinary catecholamines in chronically stressed subjects who experience sleep disturbance or in healthy subjects whose sleep was affected by nocturnal aircraft noise.

In a recent controlled study in which objective sleep testing and frequent blood sampling was employed, 24-h ACTH and cortisol plasma concentrations were significantly higher in people with insomnia than matched healthy controls. Within the 24-h period, the greatest elevations were observed in the evening and first half of the night. Also, people with insomnia with a high degree of objective sleep disturbance (sleep time $<70\%$) compared to those with a low degree of sleep disturbance secreted a higher amount of cortisol. Pulsatile analysis revealed a significantly higher number of peaks per 24 h in insomniacs than in controls ($p < 0.05$), while Cosinor analysis showed no differences in the circadian pattern of ACTH or cortisol secretion between people with insomnia and controls. Thus, insomnia is associated with an overall increase of ACTH and cortisol secretion, which, however, retains a normal circadian pattern. Also, this increase relates positively to the degree of objective sleep disturbance. These findings are consistent with a disorder of CNS hyperarousal rather than one of sleep loss, which is usually associated with no change or a decrease in cortisol secretion, or a circadian disturbance. Two recent studies found that cortisol was elevated in primary insomnia with objectively documented sleep disturbance during the evening and nighttime. However, another study found that people with insomnia without evidence of objective sleep disturbance did not report differences from controls. These findings suggest that objective sleep disturbance may be a useful index of the biological severity of the disorder.

Chronic activation or dysregulation of the stress system may play a significant role in the poor mental and physical health associated with chronic insomnia. Insomnia is associated with a higher risk of development of psychiatric disorders, such as major depression and anxiety disorders. Also, people with insomnia tend to have more health complaints than normal controls, both as children and adults, and they do appear to have poorer health in other respects. Psychosomatic-type illnesses, such as allergy, asthma, colitis, hypertension, migraine headaches, and ulcers, are reported by about twofold more people with insomnia than controls. Consistent with the above finding, a recent study that used a large general random sample of men and women (1754 people) reported that insomnia is independently associated with medical disorders such as hypertension. Also, nonspecific physical symptoms, such as headache, diarrhea, constipation, stomach discomfort, palpitations, shortness of breath, nonspecific pain, tiredness, and weakness, are also more common in people with insomnia. It can be hypothesized that these physical complaints, including those labeled as psychosomatic, reflect the chronic effects of an activation/dysregulation of the stress system in various body organs and functions. Furthermore, these data suggest that the therapeutic goal in insomnia should not be just to improve the quality or quantity of nighttime sleep. Rather, they suggest that the common practice of prescribing hypnotics only for patients with chronic insomnia at most is of limited efficacy. Sedative antidepressants alone or in combination with the newer antidepressants, e.g., selective serotonin reuptake inhibitors (SSRIs), appear to have a normalizing effect on measures, i.e., cortisol. Furthermore, the focus of psychotherapeutic and behavioral modalities, including sleep hygiene measures, should not be to just improve the emotional and physiological state of the people with insomnia

pre- or during sleep, but rather to decrease the overall emotional and physiologic hyperarousal and its underlying factors, present throughout the 24 h sleep–wake period. Finally, objective measurements of sleep using simpler methods such as actigraphy may prove useful in predicting the severity of the disorders and, therefore, the urgency for medical intervention.

Disorders of excessive daytime sleepiness and stress system Obstructive sleep apnea, the most common sleep disorder associated with excessive daytime sleepiness and fatigue, is associated with nocturnal hypoxia and sleep fragmentation. The latter conditions should be expected to be associated with an activation of the stress system. Indeed, it has been shown that urinary catecholamines, as well as plasma catecholamines measured during the nighttime, are elevated in people with sleep apnea compared to controls (Table 1). Also, using microneurography, it has been shown that obstructive apneas are associated with a surge of sympathetic nerve activity. It has been proposed that sympathetic activation in sleep apnea is one of the mechanisms leading to the development of hypertension, a condition commonly associated with sleep apnea. With similar logic, it was expected that sleep apnea would be associated with an activation of the HPA axis also. However, the few studies that have assessed the plasma cortisol levels in people with sleep apnea have failed to show any differences from controls. Also, no differences were reported in the excretion of cortisol following the abrupt withdrawal of CPAP, a definitive treatment for sleep apnea. However, whether these results represent an adaptation of the HPA axis to a chronic physical stress or are a result of incomplete assessment of the HPA axis, remains open.

Narcolepsy and idiopathic hypersomnia are two primary sleep disorders of bipolar disorders currently treated with medication such as stimulants and modafinil. The association of the two other organic disorders of excessive daytime sleepiness, narcolepsy and idiopathic hypersomnia, with the stress system has not been assessed. Preliminary data from a study that assessed the responsiveness of plasma cortisol and ACTH to the exogenous administration of CRH in individuals with idiopathic hypersomnia indicated reduced 24-h UFC levels compared to controls while ACTH response to CRH tended to be significantly higher in the patients compared to controls (Table 1). These preliminary findings suggest a subtle hypocortisolism and inferred central CRH deficiency in patients with idiopathic hypersomnia. A central CRH deficiency in these patients is consistent with their clinical profile of increased daytime sleepiness and deep nocturnal sleep (generalized hypoarousal).

The stress related cytokines (TNF-α, IL-1, and IL-6) have been suggested as potential mediators of excessive daytime sleepiness (Table 1). In 1997, a report on the morning plasma levels of TNF-α, IL-1β, and IL-6 in patients with sleep apnea, narcolepsy and idiopathic hypersomnia showed that TNF-α was significantly elevated in people with sleep apnea and narcolepsy compared to that in normal controls. Plasma IL-1β concentrations were not different between the sleep disorder patients and controls, whereas IL-6 was markedly and significantly elevated in patients with sleep apnea compared to that in normal controls. The primary factor in the elevation of plasma TNF-α values was the degree of nocturnal sleep disturbance, whereas the primary determinant for IL-6 levels was body mass index. These findings suggest that TNF-α and IL-6 might play a significant role in mediating sleepiness and fatigue in disorders of excessive daytime sleepiness in humans. Also, it was hypothesized that IL-6 plays a role in mediating sleepiness in obese patients, even those without obstructive sleep apnea, who experience, as confirmed also in the sleep laboratory, a significant degree of daytime sleepiness and fatigue. In 1996, another study reported that the circadian rhythm of TNF-α was significantly altered in patients with sleep apnea and that the peak concentration that occurred during the night in normal control subjects was not present in patients. Rather, the patients exhibited increased TNF-α concentrations in the afternoon, the time period during which concentrations in normal control subjects were at a minimum. Furthermore, in this study, it was reported that the alteration in the cytokine profiles were not normalized by the CPAP therapy. Since then, several studies have confirmed that TNF-α and IL-6 are elevated in disorders of excessive daytime sleepiness, i.e., sleep apnea and narcolepsy independently of confounding variables such as obesity. Interestingly, a recent study reported that neutralizing TNF-α is associated with a significant decrease of sleepiness in patients with sleep apnea.

Conclusion

In conclusion, the collective findings on the association between sleep and stress system appear to suggest that sleep is a major antistress mechanism that is vital in the preservation of the physical and mental homeostasis of the human species. Also, sleep loss appears to be a physical and mental stress to humans, whereas sleep disorders/disturbances may lead to a significant dysregulation of the stress system. More knowledge

on the association between sleep, sleep disorders, and stress system would lead to a better understanding of the mechanisms that underlie sleep's capacities for physical and mental restoration and to improve treatment of sleep disorders/disturbances.

Further Reading

Chrousos, G. P. (1995). The hypothalamic-pituitary-adrenal axis and immune-mediated inflammation. *New England Journal of Medicine* **332**, 1351–1362.

Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders. *JAMA* **267**, 1244–1252.

Edinger, J. D. and Means, M. K. (2005). Cognitive-behavioral therapy for primary insomnia. *Clinical Psychology Review* **25**, 539–558.

García-Borreguero, D., Wehr, T. A., Larrosa, O., et al. (2000). Glucocorticoid replacement is permissive for rapid eye movement sleep and sleep consolidation in patients with adrenal insufficiency. *Journal of Clinical Endocrinology and Metabolism* **85**, 4201–4206.

Kales, A. and Kales, J. (1984). *Evaluation and treatment of insomnia*. New York: Oxford University Press.

Krueger, J. M. and Obál, F. (1994). Sleep factors. In: Saunders, N. A. & Sullivan, C. E. (eds.) *Sleep and breathing*, pp. 79–112. New York: Dekker.

Opp, M. R. (1995). Corticotropin-releasing hormone involvement in stressor-induced alterations in sleep and in the regulation of waking. *Immunology* **5**, 127–143.

Rechtschaffen, A. and Kales, A. (1968). *A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects*. NIH Report No. 204. Bethesda, MD: National Institutes of Health.

Rechtschaffen, A., Bergmann, B. M., Everson, C. A., et al. (1989). Sleep deprivation in the rat: X. integration and discussion of the findings. *Sleep* **12**, 68–87.

Rodenbeck, A., Heuther, G., Ruther, E., et al. (2002). Interactions between evening and nocturnal cortisol secretion and sleep parameters in patients with severe chronic primary insomnia. *Neuroscience Letters* **324**, 154–163.

Rodenbeck, A., Cohrs, S., Jordan, W., et al. (2003). The sleep-improving effects of doxepin are paralleled by a normalized plasma cortisol secretion in primary insomnia. A placebo-controlled, double-blind, randomized, cross-over study followed by an open treatment over 3 weeks. *Psychopharmacology* **170**, 423–428.

Somers, V. K., Dyken, M. E., Mark, A. L., et al. (1993). Sympathetic-nerve activity during sleep in normal subjects. *New England Journal of Medicine* **328**, 303–307.

Steiger, A. and Holsboer, F. (1997). Neuropeptides and human sleep. *Sleep* **20**, 1038–1052.

Steiger, A., Antonijevic, I. A., Bohlhalter, S., et al. (1998). Effects of hormones on sleep. *Hormone Research* **49**, 125–130.

Van Cauter, E. (2005). Endocrine physiology. In: Kryger, M. H., Roth, T. & Dement, W. C. (eds.) *Principles and practices of sleep medicine* (4th ed., pp. 266–282). Philadelphia, PA: Elsevier Saunders.

VanCauter, E., Leproult, R. and Plat, L. (2000). Age-related changes in slow wave sleep and REM sleep and relationship with growth hormone and cortisol levels in healthy men. *JAMA* **284**, 861–868.

Vgontzas, A. N., Papanicolaou, D. A., Bixler, E. O., et al. (1997). Elevation of plasma cytokines in disorders of excessive daytime sleepiness: role of sleep disturbance and obesity. *Journal of Clinical Endocrinology and Metabolism* **82**, 1313–1316.

Vgontzas, A. N. and Kales, A. (1999). Sleep and its disorders. *Annual Review of Medicine* **50**, 387–400.

Vgontzas, A. N., Bixler, E. O., Lin, H.-M., et al. (2001). Chronic insomnia is associated with nyctohemeral activation of the hypothalamic-pituitary-adrenal axis: clinical implications. *Journal of Clinical Endocrinology and Metabolism* **86**, 3787–3794.

Vgontzas, A. N., Zoumakis, M., Bixler, E. O., et al. (2003). Impaired nighttime sleep in healthy old vs. young adults is associated with elevated plasma IL-6 and cortisol levels: physiologic and therapeutic implications. *Journal of Clinical Endocrinology and Metabolism* **88**, 2087–2095.

Vgontzas, A. N., Zoumakis, M., Bixler, E. O., et al. (2004). Adverse effects of modest sleep restriction on sleepiness, performance, and inflammatory cytokines. *Journal of Clinical Endocrinology and Metabolism* **89**, 2119–2126.

Smoking and Stress

F J McClernon
Duke University Medical Center, Durham, NC, USA
D G Gilbert
Southern Illinois University-Carbondale, Carbondale, IL, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D G Gilbert and F J McClernon, volume 3, pp 458–465, © 2000, Elsevier Inc.

Introduction
Personality Traits, Genetic, and Biological Predictors of Smoking and Stress
Smoking as a Coping Response
Smoking Cessation, Stress, and Coping
The Psychobiology of Smoking, Coping, and Stress Diasthesis
Summary

Glossary

Acetylcholine	A neurotransmitter in the brain that has receptors on nerve cells that nicotine stimulates and promotes the firing of.
Affect	Moods, emotions, and feelings, including negative affect (e.g., anxiety, depressive mood, and anger) and positive affect (e.g., euphoria, vigor, and pleasant feelings).
Cognitive dysfunction	The impairment or biasing of information processes such as perceptual, attentional, and memorial functions that can be caused by temporary physiological and mood states, as well as by longer-lasting forms of psychopathology, neurological abnormality, or temperamental bias.
Dopamine	A neurotransmitter in the brain having a number of functions, including the enhancement of attention and signaling of reward in response to appetitive stimuli.
Neurotransmitter	A chemical released at the synapse (junction) between nerves that promotes the conduction of nerves impulses from one nerve to another.
Nicotine	The primary psychoactive component of tobacco smoke; it binds to endogenous nicotinic acetylcholine receptors, which results in a variety of neuronal actions.
Situation x trait adaptive response (STAR) model	An integrative model that emphasizes the interaction of situational factors with genetic and personality factors in determining who is likely to become a smoker, how rewarding smoking will be, and when and how often smoking is likely to be rewarding to the individual as a coping tool.
Tar	The mass of some 5000 components of tobacco smoke, excluding nicotine and gases (e.g., carbon monoxide).

Introduction

Smoking remains the number one preventable cause of death and disability in the United States and is a burden on people, economies, and health-care systems worldwide. Knowledge about the relation of smoking and stress has expanded exponentially in the last 3 decades as a result of new findings in the fields of genetics, neuroscience, and clinical research. This research has revolutionized thinking about who smokes and why it is so difficult for many to quit. Genes have been found to play a significant role in determining who becomes and remains a smoker. Some of the genes that predispose some people to smoke predispose the same individuals to experience higher than normal levels of stress and negative mood states. These and related findings indicate that there are many clinically and theoretically important relations among smoking, stress, and coping.

Personality Traits, Genetic, and Biological Predictors of Smoking and Stress

Evidence shows clear and substantial relations among smoking, stress-related personality traits, and psychopathology. Further, all these characteristics are largely heritable (passed on from generation to generation via genes) and share common genetic bases.

Stress, Personality, Psychopathology, and Smoking

Individuals with specific psychiatric problems and personality traits leading to frequent or intense personal distress are substantially more likely to smoke. Smoking rates are disproportionately high in individuals with psychiatric conditions, including major depressive disorder, various anxiety disorders, schizophrenia,

antisocial personality disorder, attention deficit hyperactivity disorder (ADHD), and various substance abuse disorders, including alcoholism. The prevalence of smoking in individuals with a lifetime history of these conditions is two to three times higher than in unaffected individuals. The personality traits associated with the increased prevalence of smoking reflect dispositions that are similar in nature to psychiatric disorders. For example, the personality trait of neuroticism (the general tendency to experience negative affect and stress) has been found to be associated with higher prevalence of smoking. Smoking (and drug abuse more broadly) is also much more prevalent among individuals high in impulsiveness or sensation-seeking traits. In addition to finding relations between these traits and smoking among adults, longitudinal studies have shown that children and adolescents with these traits are more likely to smoke as adults (as well as to use other drugs). Taken together, these data suggest that smoking behavior is not randomly distributed in the adult population and also support the notion that nicotine plays some role in the psychological life of the smoker beyond simple addiction or withdrawal alleviation.

Identical Genes Related to Smoking, Psychopathology, and Personality Traits

Results from family, twin, and molecular genetic studies demonstrate that genetic factors play an important role in smoking behavior and stress responses. Genes account for approximately half of the variability in who will smoke and who will not. Further, many of the same genes that predispose individuals to smoke are also responsible, in part, for predisposing individuals to the stress-related personality traits and psychiatric disorders previously mentioned. For instance, the same genes that contribute to a vulnerability to neuroticism also appear to dispose individuals to clinical depression and to smoking. Similarly, the same genes that contribute to impulsivity and sensation seeking also contribute to a variety of compulsive behaviors, including drug abuse, antisocial behavior, and smoking.

The identification of some of the many genes that contribute to smoking and smoking-related traits has been one of the major accomplishments of recent years. Genetically influenced biological differences between people in (1) brain neurotransmitter receptors and (2) liver enzymes related to nicotine metabolism appear to play roles in determining smoking. Some of the most exciting insights in this area are related to the neurotransmitter dopamine. Dopamine is intimately associated with coping, stress, natural reward, and the pleasurable effects of abused drugs, including nicotine. A number of genes combine to determine a person's brain dopamine system functioning. Those individuals with low dopamineric system functioning levels are more likely than others to smoke, abuse alcohol and other drugs, and have impulsive and sensation-seeking personality traits. Thus, these dopamine genes appear to contribute to the association of smoking with these personality traits and substance-abuse patterns.

Smoking as a Coping Response

These personality traits and psychiatric conditions are associated with distressing cognitive and emotional states. To the degree smoking is effective at ameliorating these undesirable cognitive or affective states, coping by means of smoking will occur and will be experienced as rewarding by a smoker. It is not surprising, then, that those individuals who frequently experience cognitive and affective states that can be improved by smoking are more likely than others to find smoking to be a useful and reinforcing tool (and thus smoke more often).

Although smokers as a group report smoking to reduce negative affect; cope with stress; reduce weight or appetite; and enhance stimulation, cognitive performance, and pleasure, the relative importance of these smoking motivations vary across (and also within) smokers. Consistent with the coping model, those who are depression prone are more likely to say they smoke to reduce depressive moods and thoughts, whereas those with attentional problems say they smoke more often to enhance attention and, in turn, those who are prone to anxiety are more likely to say they smoke to alleviate anxiety. Similarly, individuals who are motivated to keep from gaining weight value the fact that smoking helps maintain a lower body weight. Thus, the motivations to smoke are to some extent consistent with personality traits of the smokers. This tendency to smoke for reasons related to personality traits is called trait-adaptive responding because smoking reflects an adaptive coping response related to the smoker's specific personality traits (see **Figure 1**).

Smoking/Nicotine as a Coping Response

Withdrawal alleviation As previously suggested, one theory proposes that smoking alleviates stress only to the extent that nicotine withdrawal is stressful to smokers. These withdrawal alleviation theories are supported by the simple fact that smokers find abstinence from smoking stressful and smoking alleviates this stress. A related theory suggests that smokers so come to associate withdrawal-induced stress with negative states that any negative state – whether during withdrawal or not – cues smoking behavior.

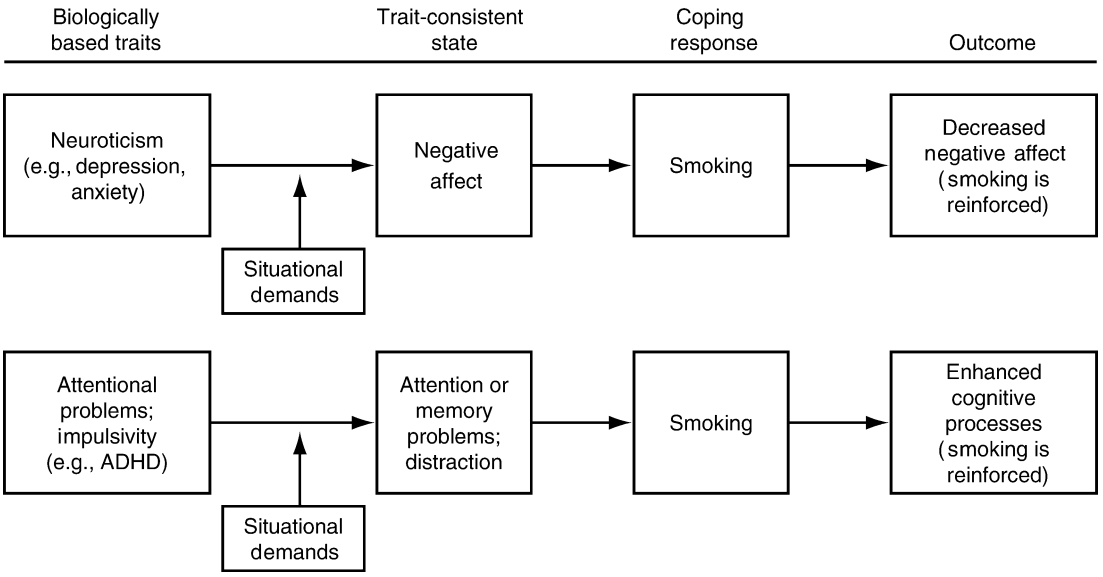


Figure 1 Biologically based traits increase the frequency and severity of trait-consistent states. To the extent that these states are improved by smoking/nicotine, smoking is reinforced.

Although withdrawal alleviation models of smoking and stress are compelling, they cannot provide a complete picture of nicotine dependence. The withdrawal alleviation model suggests the stress of withdrawal is temporary; once withdrawal subsides the now-ex-smokers will return to their presmoking affective and cognitive states. Indeed, many studies have reported temporary increases in negative affect, stress, and difficulty concentrating following quitting that return to or improve beyond prequit levels in a matter of weeks. However, these studies have been limited by a number of factors that might preclude accurate characterization of withdrawal in smokers. First, because these studies do not take into account data of smokers who relapse (i.e., smokers for whom quitting is most difficult), the findings minimize the negative affect and distress the majority of smokers experience on quitting. Second, these studies have often not included a continuing-to-smoke control group, so other factors affecting mood and cognition, such as the passage of time, are also not accounted for.

When these two limitations are eliminated from research designs, a different picture of withdrawal emerges. Recent longitudinal studies that minimize relapse and include adequate control groups have observed transient increases in negative affect, stress, and difficulty concentrating that do not return to baseline levels, or the approximate levels of these symptoms in a continuing-to-smoke control group. These findings suggest that smokers may smoke, in part, to reduce underlying distressful cognitive and affective states that were present prior to becoming a smoker, which, on quitting smoking, are rerevealed. Thus, smokers may smoke in order to alleviate withdrawal symptoms to a degree, but also smoke in order to self-medicate underlying trait-based problems. It should be noted, however, that well-designed studies to date have accurately assessed the effects of quitting smoking on negative affect for only 45 days of abstinence. Thus, it is possible that after a longer period of time effects of quitting do resolve in most quitters.

Absolute inherent effects nicotine A growing body of research suggests that nicotine can, under specific situations, improve mood and cognitive performance, even in smokers who are only minimally deprived of nicotine or in nonsmokers. Laboratory studies show that smoking and other forms of nicotine delivery usually reduce negative moods and emotions in situations in which stressors are ambiguous and/or not dominant, such as situations in which an individual is anticipating experiencing stress. In contrast, when the context includes clear, immediate, and potent threat stimuli, negative moods are not impacted by smoking or nicotine.

Many experts and some experimental evidence suggest that nicotine can also indirectly promote positive affect and minimize negative affect in situations in which nicotine contributes to goal achievement. For example, when they need to concentrate for a long period of time to accomplish a goal, such as writing a paper, smokers say they smoke to sustain their concentration and thereby achieve their goal. Laboratory studies have also shown nicotine to facilitate cognitive functions, including vigilance and memory,

in both habitual tobacco users and in individuals with no prior exposure to nicotine. In addition, other studies have shown that nicotine may improve affect and reduce stress by helping smokers focus on emotionally positive or neutral distractors. Research suggests that nicotine is especially effective at reducing stress when there is a substantial degree of choice as to whether or not to attend to stressful cues and when some of these nonstressful cues are emotionally positive.

A two-factor model of smoking A combined two-factor model proposes that smokers smoke in order to (1) prevent or alleviate withdrawal and (2) gain beneficial effects inherent in nicotine. Among heavy smokers, the acute effects of abstinence resolve across the first week or two of abstinence; thereafter, underlying psychological and pathological traits are rerevealed and greater stress, negative affect, and difficulty concentrating are experienced. Similarly, the attention and mood-enhancing effects of most habitually used stimulant drugs (caffeine, amphetamine, and cocaine) appear to reflect the combined influence of such inherent and withdrawal-alleviation mechanisms.

Habit, sensory aspects, and other factors Smokers may smoke for numerous reasons other than those already described. Many smokers find the sensory aspects of smoking pleasurable (e.g., the feel of smoke in lungs and the taste), and these sensory cues further reinforce smoking behavior. Further, many smokers, regardless of why they started smoking in the first place, may continue to smoke, largely out of habit. Finally, recent evidence suggests that nonnicotine chemicals in tobacco smoke may have anti-depressant qualities that may further ameliorate negative affect and reinforce smoking behavior.

Smoking Cessation, Stress, and Coping

Based on the empirical evidence showing that smoking cessation is frequently a stressful process resulting in more coping failures than successes, clinicians and researchers have spent much time and energy on minimizing smoking-cessation-related stress and maximizing coping.

Psychological and Behavioral Therapies for Smoking Cessation

Most of the widely used smoking cessation programs include several hours of skills training in cognitive and behavioral techniques designed to help the smoker cope not only with the cravings to smoke but also with negative mood states frequently associated with quitting and known to promote relapse to smoking. Such techniques include behavioral relaxation training and learning to think about stressful situations in a manner that is less likely to produce negative emotional states. Currently popular pharmacological treatments, such as nicotine gum, nicotine patches, and the antidepressant drug Zyban, also include a complex therapeutic package that includes behavioral relaxation and coping components, along with information about the cessation process and social support.

Antidepressant Medication Therapies for Smoking Cessation

Consistent with the view that smoking motivation and the effects of smoking cessation are mediated by brain mechanisms related to negative affect and depression, one antidepressant medication, Zyban, has been found to be beneficial in helping individuals maintain smoking abstinence. This U.S. Food and Drug Administration (FDA)-approved smoking-cessation drug is also marketed under a different brand name as an antidepressant medication.

The Psychobiology of Smoking, Coping, and Stress Diasthesis

During withdrawal from nicotine, smokers with depressive tendencies have been found to exhibit brain wave (via electroencephalogram, EEG) asymmetry profiles characteristic of depressed individuals. These abstinence-induced EEG asymmetries appear to last for at least a month after quitting. The EEG asymmetry seen in these individuals is a relatively greater EEG slowing of the left (compared to the right) frontal cortex subsequent to quitting. Left-hemisphere processing is associated with active coping, verbal behavior, and positive affect, whereas right-hemisphere processing is more associated with negative mood states and withdrawal. These and related findings suggest that nicotine may act as an antidepressant in part because in individuals prone to depression it tends to increase the activation more in the left than in the right frontal cortex. Overall, nicotine tends to increase brain activation as measured by EEG, whereas smoking abstinence tends to decrease EEG activity in both hemispheres, although to a greater extent in the left relative than in the right frontal cortex. This decrease in EEG activation corresponds to the sustained decrease in ability to concentrate also found to last for at least a month after quitting smoking.

The EEG-activating effects of smoking are thought to reflect the effects of nicotine on brain receptors for the neurotransmitter acetylcholine. Nicotine has the ability to stimulate cholinergic receptors throughout the body. The stimulation of these cholinergic receptors can result in a cascade of events, including increases in heart rate, blood pressure, blood cortisol concentrations and the activation of other neurotransmitter systems in the brain. Another important brain system influenced by nicotine uses dopamine as its neurotransmitter. Dopamine (DA) is associated with brain regions involved in attention, emotion, coping, and reward. DA is involved in coping in both positive and negative emotional states. Nicotine-induced increments in brain mesolimbic DA activity may increase the ability of the organism to avoid aversive stimuli and to obtain positive ones.

The effects of smoking on EEG, mood, and attention discussed so far have been found to be due primarily to the effects of nicotine; however, recent evidence suggests that an unidentified component of tobacco smoke (other than nicotine) inhibits monoamine oxidase (MAO) activity in the brain in the same manner as some antidepressant medications. Thus, it may be that some of smoking's antidepressant, mood-enhancing, and cognitive benefits come from the inhibition of MAO activity by something other than nicotine.

Summary

Environment, genetics, personality, and psychopathologic traits interact to determine smoker status and the type of smoker an individual becomes. Individuals genetically and otherwise predisposed to negative affect, sensation seeking, and/or suboptimal cognitive performance are more likely to be exposed to smoking environments, start smoking in order to seek negative affect relief or as an act of sensation seeking, and find smoking reinforcing. The widely held belief among smokers and nonsmokers that smoking reduces negative affect and enhances cognition leads predisposed individuals to experiment with smoking as a form of self-medication, to minimize negative affect, and to maximize cognitive functioning. The affect-modulating and cognitive-enhancing effects of nicotine are more reinforcing to those who need and value them the most – those in stressful circumstances and those with a vulnerability to cognitive and affective dysfunction. When individuals who are predisposed to smoking try to quit, they lose the beneficial effects of nicotine and are especially vulnerable to experiencing the cognitive or affective state to which they are predisposed. Thus, depression-prone individuals are more likely to become depressed, anxiety-prone individuals are more prone to anxiety, and attentionally disordered quitters are disposed to suffer more severe attention problems. The enhanced severity of cognitive and affective responses to quitting contributes to the failure of vulnerable individuals to maintain abstinence after efforts to quit smoking. Smoking in individuals who are vulnerable to smoking can thus be seen as an attempt to adapt to their situation by altering their cognitive and/or affective state.

However, affective and cognitive traits also alter the probability of smoking by more indirect mechanisms. Both affective and cognitive traits influence long-term goals by a number of indirect mechanisms. For example, risk-taking and other reward and sensation-seeking behaviors make certain individuals temperamentally insensitive to cues of potential punishment (e.g., health warnings about smoking) and promote cognitive, social, and behavioral sequences that reinforce smoking and other risk-taking behavior. The sensation-seeking life style, which includes smoking, drinking, and a focus on pleasure, may be so reinforcing to individuals with a strong genetic disposition that their cognitive motivational schemas include no plans to quit any risky behaviors. In contrast, negative-affect-prone individuals are likely to focus on the immediate relief of negative affect produced by smoking rather than on the long-term health consequences. Individuals who are more psychobiologically stable will see less benefit from smoking and be less inclined to take up or continue smoking because the perceived risks outweigh the perceived gains.

See Also the Following Articles

Alcohol and Stress: Social and Psychological Aspects; Alcohol, Alcoholism and Stress: A Psychobiological Perspective; Dopamine, Central; Drug Use and Abuse; Health Behavior and Stress; Comorbid Disorders and Stress.

Further Reading

Comings, D. E., Ferry, L., Bradshaw-Robinson, S., et al. (1996). The dopamine D2 receptor (DRD2) gene: a genetic risk factor in smoking. *Pharmacogenetics* **6**, 73–79.

Gilbert, D. G. (1995). *Smoking: individual differences, psychopathology, and emotion*. Washington, DC: Taylor and Francis.

Gilbert, D. G. (1997). The situation x trait adaptive response (STAR) model of substance use and craving. *Human Psychopharmacology: Clinical and Experimental* **12**, S89–S102.

Gilbert, D. G., McClernon, F. J., Rabinovich, N. E., et al. (1998). Effects of smoking abstinence on mood and craving in men: influences of negative-affect-related personality traits, habitual nicotine intake, and repeated measurements. *Personality and Individual Differences* **25**, 399–423.

Gilbert, D. G., McClernon, F. J., Rabinovich, N. E., et al. (2002). Mood disturbance fails to resolve across 31 days of cigarette abstinence in women. *Journal of Consulting & Clinical Psychology* **70**, 142–152.

Gilbert, D. G., McClernon, F. J., Rabinovich, N. E., et al. (2004). Effects of quitting smoking on EEG activation and attention last for more than 31 days and are more severe with stress, dependence, DRD2 A1 allele, and depressive traits. *Nicotine & Tobacco Research* **6**, 249–267.

Kandel, D. B. and Davies, M. (1986). Adult sequelae of adolescent depressive symptoms. *Archives of General Psychiatry* **43**, 255–262.

Kassel, J. D. (1997). Smoking and attention: a review and reformulation of the stimulus-filter hypothesis. *Clinical Psychology Review* **17**, 451–478.

Kassel, J. D., Stroud, L. R. and Paronis, C. A. (2003). Smoking, stress, and negative affect: correlation, causation, and context across stages of smoking. *Psychological Bulletin* **129**, 270–304.

Kendler, K. S., Neale, M. C., MacLean, C. J., et al. (1993). Smoking and major depression: a causal analysis. *Archives of General Psychiatry* **50**, 36–43.

Levin, E. D., Conners, C. K., Sparrow, E., et al. (1996). Nicotine effects on adults with attention-deficit/hyperactivity disorder. *Psychopharmacology* **123**, 55–63.

Scarr, S. and McCartney, K. (1983). How people make their own environments: a theory of genotype x environment effects. *Child Development* **54**, 424–435.

Shiffman, S. A. (1986). Cluster analytic classification of smoking relapse episodes. *Addictive Behaviors* **11**, 295–307.

Social Capital

I Kawachi
Harvard School of Public Health, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by I Kawachi, volume 3, pp 466–467, © 2000, Elsevier Inc.

Forms and Functions of Social Capital
Measurement of Social Capital
Social Capital and Population Health
Mechanisms Linking Social Capital and Health
Negative Effects of Social Capital

Glossary

Bonding social capital	Social connections that exist within homogeneous groups, for example the strong ties that connect family members, close friends, or neighbors who are alike with respect to race/ethnicity, social class, and religion.
Bridging social capital	Social connections that cross race/ethnic, class, and other boundaries.
Social capital	The resources available to individuals and groups through their social connections and membership in community networks.

Forms and Functions of Social Capital

Social capital is the resources available to individuals through their social connections and membership in community networks. In contrast to financial capital, which resides in people's bank accounts, or human capital, which is embodied in individuals' investments in education and job training, social capital inheres in the structure and quality of social relationships between individuals.

The concept was originally developed in sociology and political science. According to James Coleman, social capital can take on several different forms, such as levels of trust within a social structure, norms and sanctions, information channels, and appropriable social organizations. An example of an appropriable social organization is the case of a residents' association in an urban housing project that formed initially for the purpose of pressuring builders to fix various problems (e.g., leaks and crumbling sidewalks). After the original problems were solved, the organization remained as available social capital within the community to improve the quality of life of residents in other ways.

In fields outside of health, social capital has been applied in a variety of settings ranging from the smooth and effective functioning of government institutions to the ability of communities to solve the

problems of collective action, such as intervening to prevent the occurrence of juvenile delinquency. More recently, empirical studies have begun to link social capital to population health outcomes.

Measurement of Social Capital

Two approaches have been used to measure social capital: (1) surveys and (2) direct social observations. In the survey-based approach, researchers inquired about a range of attitudes (trust of others and expectations of reciprocity and mutual aid) as well as behaviors (e.g., membership in community organizations, levels of voluntarism, and intensity of informal social interactions with neighbors) that are thought to be indicative of social capital. By contrast, direct social observation may involve the use of experimental methods, such as the letter-drop experiment in which stamped addressed envelopes are deliberately dropped on street corners to determine the proportion of letters that are subsequently mailed to the addressee by strangers (an experimental indicator of reciprocity). More commonly, direct observational approaches emphasize the observable characteristics of communities that may serve as proxies for the extent of social capital, such as the density of voluntary associations within a given locality.

Given the well-established tradition in community psychology of studying the characteristics of cohesive communities, there is legitimate debate over the extent to which the concept of social capital represents old wine in new bottles. In community psychology, concepts such as the psychological sense of community, community competence, and neighboring all predate the current vogue for the term social capital. In addition, there is an extensive prior literature in health psychology and social epidemiology on the related constructs of social support and social integration. However, an important distinction between social support and social capital is that, whereas the former variable is measured at the individual level, the latter is most often conceptualized and measured at the community or area level. That is, social capital has most often been considered to be a property of the collective (neighborhoods, cities, or even regions).

If social capital is indeed a collective property, a critical question is whether aggregated responses to social surveys (e.g., in response to questions about trusting neighbors) genuinely represent a group characteristic or whether they are confounded by other characteristics of the individual respondents that correlate with perceptions of social capital. Some evidence suggests that aggregated survey responses actually capture genuine place-based differences. An analysis of the Community Survey of the Project on Human Development in Chicago Neighborhoods reported that even after controlling for the sociodemographic attributes of individual respondents (age, sex, race/ethnic status, income, and schooling), there were significant between-neighborhood variations in levels of social capital as measured by perceptions of trust.

Finally, research has increasingly pointed out the important distinction between bonding and bridging social capital. Bonding social capital refers to the social connections that exist within homogeneous groups, for example the strong ties that connect family members, close friends, or neighbors who are alike with respect to race/ethnicity, social class, and religion. By contrast, bridging social capital refers to social connections that cross race/ethnic, class, and other boundaries.

Social Capital and Population Health

Empirical studies have linked social capital to a growing number of population health outcomes, including mortality, mental health, health-related behaviors, self-rated health, and disability. In one of the first ecological analyses, indicators of social capital at the U.S. state level (levels of trust, perceptions of reciprocity, and density of civic associations) were strongly correlated with all-cause and cause-specific mortality rates. Subsequent investigations extended these findings down to the neighborhood level. In a study based in the 343 neighborhoods of Chicago, the levels of trust, reciprocity, and group membership were found to be significantly inversely correlated with mortality rates for residents, even after controlling for the degree of economic deprivation.

Increasingly, studies of social capital and health have used multilevel analysis, with interesting results. In one study of mental health among African American residents of a disadvantaged community in Alabama, higher bonding social capital was found to be associated with worse mental health outcomes, whereas bridging social capital was found to be associated with better mental health. In other words, within the context of a highly disadvantaged community, a greater level of social connection and involvement may be associated with worse health outcomes because of the strains associated with the obligation to help others in such settings. In another study analyzing self-rated health within the 2000 Social Capital Community Benchmark Survey (conducted in 40 U.S. communities), a complex interaction was found between community levels of social trust and individual levels of trust. Among individuals who expressed a high level of trust of others, there was a positive influence (on self-rated health) of living in

communities with higher levels of trust. On the other hand, the opposite was found for individuals expressing mistrust of others in their community. For these individuals, the higher the community's level of trust, the worse they felt.

Mechanisms Linking Social Capital and Health

Several mechanisms have been hypothesized through which social capital can influence health. They include:

1. The ability of a community to undertake collective action (e.g., organizing to pass local ordnances that restrict smoking in public places or lobbying to fight the closure of local clinics).
2. The ability of a community to enforce healthy norms by exercising informal social control over deviant behaviors (e.g., underage drinking, smoking, and drug abuse).
3. The faster diffusion of innovation through information channels (e.g., new knowledge about a new cancer screening test).
4. The exchange of social support between members of the community (e.g., the provisions of emotional support and stress-buffering or the exchange of instrumental aid, such as cash loans, during emergencies).

It is important to emphasize, however, that not all kinds of social capital result in beneficial outcomes for health. Like any form of capital (e.g., financial capital), social capital can be used for pro-social as well as antisocial ends and can result in deleterious as well as beneficial health outcomes.

Negative Effects of Social Capital

Social capital is not a panacea for health promotion. Strong social ties within a group or a community (high bonding social capital) can imply a high level of social exchange, which in turn translates into role strain for certain members of the in-group (the support providers, who are often women). Strong bonding social capital can often coexist with conflict between the group and outside social groups low bridging social capital). In some instances, high social capital within a community has been channeled to exclude outsiders from entering the community – for example, by enforcing residential segregation. Strong levels of social cohesion can also imply restrictions on the freedom of individual expression as well as a downleveling of achievement norms. In addition, some forms of social capital, such as criminal gangs, may provide benefits to its members, but do very little to foster social cohesion with the rest of society. Finally, some have criticized the concept of social capital for encouraging the blaming of victims (or, in this case, the blaming of communities) for their problems. These criticisms emphasize the caution warranted in applying the concept of social capital in health promotion.

See Also the Following Articles

Community Studies; Social Support.

Further Reading

Coleman, J. S. (1990). *Foundations of social theory.* Cambridge, MA: Harvard University Press.

Kawachi, I., Kennedy, B. P., Lochner, K., et al. (1997). Social capital, income inequality, and mortality. *American Journal of Public Health* **87**, 1491–1498.

Kawachi, I., Kim, D. J., Coutts, A., et al. (2004). Reconciling the three accounts of social capital. *International Journal of Epidemiology* **33**(4), 682–690.

Lochner, K., Kawachi, I., Brennan, R. T., et al. (2003). Social capital and neighborhood mortality rates in Chicago. *Social Science & Medicine* **56**(8), 1797–1805.

Lochner, K., Kawachi, I. and Kennedy, B. P. (1999). Social capital: a guide to its measurement. *Health and Place* **5**, 259–270.

Mitchell, C. U. and LaGory, M. (2002). Social capital and mental distress in an impoverished community. *City & Community* **1**(2), 195–215.

Szreter, S. and Woolcock, M. (2004). Health by association?: social capital, social theory, and the political economy of public health. *International Journal of Epidemiology* **33**(4), 650–667.

Subramanian, S. V., Kim, D. J. and Kawachi, I. (2002). Social trust and self-rated health in US communities: a multilevel analysis. *Journal of Urban Health* **79**(4, supplement 1), S21–S34.

Subramanian, S. V., Lochner, K. and Kawachi, I. (2003). Neighborhood differences in social capital in the US: compositional artifact or a contextual construct? *Health and Place* **9**, 33–44.

Social Conflict *See:* Social Networks and Social Isolation; Income Levels and Stress; Social Status and Stress; Environmental Factors.

Social Networks and Social Isolation

L F Berkman
Harvard School of Public Health, Boston, MA, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction
Theoretical Orientations
Health, Social Networks, and Social Integration
A Conceptual Model Linking Social Networks to Health
Conclusion

Glossary

Appraisal support	Help in decision making, giving appropriate feedback, or help deciding which course of action to take.
Emotional support	The amount of loving and caring, sympathy and understanding, and/or esteem and value available from others.
Instrumental support	Help, aid, or assistance with tangible needs such as getting groceries, getting to appointments, and cooking; aid in kind, money, or labor.
Social networks	The web of social relationships that surrounds an individual and the characteristics of those ties.
Social participation and engagement	The enactment or interaction of social ties in real-life activity. Meeting with friends and family, attending social or religious functions, and participating in occupational or social roles or in group recreation are examples.

Introduction

Over the last 25 years, there have been dozens of articles and books on issues related to social networks and social support. It is now widely recognized that social relationships and affiliation have powerful effects on physical and mental health for a number of reasons.

When investigators write about the impact of social relationships on health, many terms are used loosely and interchangeably, including social networks, social support, social ties and social integration. The aim of this article is to (1) discuss theoretical orientations from diverse disciplines that we believe are fundamental to advancing research in this area, (2) briefly review findings related to mortality and biological pathways leading to poor health, (3) provide a set of definitions of networks and aspects of networks and support, and (4) present an overarching model that integrates multilevel phenomena.

Theoretical Orientations

There are several sets of theories that form the bedrock for the empirical investigation of social relationships and their influence on health. The earliest theories came from sociologists such as Emile Durkheim, as well as from psychoanalysts such as John Bowlby, who first formulated attachment theory. A major wave of conceptual development also came from anthropologists, including Bott, Barnes, and Mitchell, and from quantitative sociologists, such as Fischer, Laumann, Wellman, and Marsden, who, along with others, developed social network analysis. This eclectic mix of theoretical approaches coupled with the contributions of epidemiolgists form the foundation of research on social ties and health.

Durkheim's contribution to the study of the relationship between society and health is immeasurable. Perhaps most important is his contribution to the understanding of how social integration and cohesion influence suicide. Durkheim's primary aim was to explain how individual pathology was a function of social dynamics. In light of recent attention to upstream determinants of health, Durkheim's work reemerges with great relevance today.

Social Network Theory: A New Way of Looking at Social Structure and Community

During the mid-1950s, a number of British anthropologists found it increasingly difficult to understand the behavior of either individuals or groups on the basis of traditional categories such as kin groups, tribes, or villages. Barnes and Bott developed the concept social networks to analyze ties that cut across traditional kinship, residential, and class groups and to explain behaviors they observed such as access to jobs, political activity, and marital roles. The development of social network models provided a way to view the structural properties of relationships among people.

Network analysis "focuses on the characteristic patterns of ties between actors in a social system rather than on characteristics of the individual actors themselves and use these descriptions to study how these social structures constrain network member's behavior" (Hall and Wellman, 1985: 26). Network analysis focuses on the structure and composition of

the network and on the contents or specific resources that flow through those networks. The strength of social network theory rests on the testable assumption that the social structure of the network itself is largely responsible for determining individual behavior and attitudes by shaping the flow of resources that determine access to opportunities and constraints on behavior.

Health, Social Networks, and Social Integration

From the mid-1970s to the present, there have been a series of studies consistently showing that the lack of social ties or social networks predicts mortality from almost every cause of death. There are now a number of studies, relating outcomes to morbidity and biomarkers of stress or resilience to inflammation, metabolic and neuroendocrine functioning, and cardiovascular reactivity, that also support the hypothesis that social networks impact health and well-being directly through multiple pathophysiological mechanisms. Social networks and support may both directly affect health outcomes and buffer the effects of other stressful experiences.

Here, we briefly review the mortality studies. In the first of these studies from Alameda County, men and women who lacked ties to others (in this case, based on an index assessing contacts with friends and relatives, marital status, and church and group membership) were 1.9–3.1 times more likely to die in a 9-year follow-up period than those who had many more contacts. Another study, in Tecumseh, Michigan, showed a similar strength of positive association for men, but not for women, between social connectedness/social participation and mortality risk over a 10–12 year period; this study has the additional strength that it controlled for some biomedical predictors assessed from physical examination (e.g., cholesterol, blood pressure, and respiratory function).

Similar results have been reported from several studies in the United States and form three in Scandinavia. Investigators working on a study in Evans County, Georgia, found risks to be significant in older White men and women without social networks, even when biomedical and sociodemographic risk factors were controlled for, although some racial and gender differences were observed. Two studies in Sweden reported significantly increased risks among socially isolated adults, and a study of men and women in eastern Finland found that social connections were related to mortality risk for men, but not for women, independently of standard cardiovascular risk factors.

Studies of older men and women confirm the continued importance of these relationships into late life. Furthermore, two studies of large cohorts of men and women in a large health maintenance organization and 32,000 male health professionals suggested that social networks are, in general, more strongly related to mortality than to the incidence or onset of disease.

Two more recent studies in Danish men and in Japanese men and women further indicated that aspects of social isolation or social support are related to mortality. Virtually all these studies found that people who are socially isolated or disconnected from others are between two and five times more likely to die from all causes, compared to those who maintain strong ties to friends, family, and community.

Social networks and support have been found to predict a very broad array of other health outcomes from survival post-myocardial infarction to disease progression, functioning, and the onset and course of infectious diseases. Several studies suggested that social isolation is related to C-reactive protein (especially in men), measures of allostatic load, fibrinogen and cardiovascular reactivity, and neuroendocrine function. Although observational studies have been reasonably consistent in linking social isolation to outcome, psychosocial interventions have not been as successful at producing changes in social isolation or support that translate into better health outcomes.

A Conceptual Model Linking Social Networks to Health

Although the power of using measures of networks or social integration to predict health outcomes is indisputable, the interpretation of what the measures actually measure has been open to much debate. Hall and Wellman have appropriately commented that much of the work in social epidemiology has used the term social networks metaphorically because rarely have investigators conformed to more standard assessments used in network analysis. This criticism has been duly noted, and several calls have gone out to develop a second generation of network measures.

A second wave of research developed in reaction to this early work and as an outgrowth of work in health psychology, which has turned the orientation of the field in several ways. These social scientists focused on the qualitative aspects of social relations (i.e., their provision of social support or, conversely, the detrimental aspects of relationships) rather than on the elaboration of the structural aspects of social networks. Most of these investigators followed the assumption that what is most important about networks is the support functions they provide. Although

social support is among the primary pathways by which social networks may influence physical and mental health status, it is not the only critical pathway. Moreover, the exclusive study of more proximal pathways detracts from the need to focus on the social context and structural underpinnings that may significantly influence the types and extent of social support provided.

In order to have a comprehensive framework in which to explain these phenomena, it is helpful to move upstream and return to a more Durkheimian orientation to network structure and social context. It is critical to maintain a view of social networks as lodged within those larger social and cultural contexts that shape the structure of networks.

Conceptually, social networks are embedded in a macro-social environment in which large-scale social forces may influence network structure, which in turn influences a cascading causal process beginning with the macro-social to psychobiological processes to impact health. In this framework, social networks are embedded in a larger social and cultural context in which upstream forces are seen to condition network structure.

Networks may operate at the behavioral level through at least five primary pathways: (1) provision of social support, (2) social influence, (3) social engagement and attachment, (4) person-to-person contact, and (5) access to resources and material goods. These psychosocial and behavioral processes may influence even more proximate pathways to health status, including direct physiological responses; psychological states, including self-esteem, self-efficacy, and depression; health-damaging behaviors such as tobacco consumption or high-risk sexual activity, health-promoting behavior such as appropriate health service use and exercise, and exposure to infectious disease agents such as human immunodeficiency virus (HIV), other sexually transmitted diseases (STDs), and tuberculosis.

The Assessment of Social Networks

Social networks might be defined as the web of social relationships that surround an individual and the characteristics of those ties. Network characteristics are:

- Range or size: the number of network members.
- Density: the extent to which the members are connected to each other.
- Boundedness: the degree to which they are defined on the basis of traditional structures such as kin, work, and neighborhood.
- Homogeneity: the extent to which individuals are similar to each other in a network.

Related to network structure, the characteristics of individual ties include:

- Frequency of contact.
- Multiplexity: the number of types of transactions or support flowing through ties.
- Duration: the length of time an individual knows another
- Reciprocity: the extent to which exchanges are reciprocal.

Downstream Social and Behavioral Pathways

We next briefly review the primary pathways by which networks may influence health.

Social support Moving downstream, we now consider the mediating pathways by which networks might influence health status. Most obviously, the structure of network ties influences health via the provision of many kinds of support. This framework immediately acknowledges that not all ties are supportive and that there is variation in the type, frequency, intensity, and extent of support provided. For example, some ties provide several types of support, whereas other ties are specialized and provide only one type. Social support is typically divided into emotional, instrumental, appraisal, and informational support.

Emotional support is related to the amount of "love and caring, sympathy and understanding and/or esteem or value available from others" (Thoits 1995: 64). Emotional support is most often provided by a confidant or intimate other, although less intimate ties can provide such support under circumscribed conditions. Instrumental support refers to help, aid, or assistance with tangible needs such as getting groceries, getting to appointments, phoning, cooking, cleaning, and paying bills. House identifies instrumental support as aid in kind, money, or labor. Appraisal support is help in decision making, help by giving appropriate feedback, or help in deciding which course of action to take. Informational support is the provision of advice or information in the service of particular needs. Emotional, appraisal, and informational support are often difficult to disaggregate and have various other definitions (e.g., self-esteem support).

Perhaps even deeper than support are the ways in which social relationships provide a basis for intimacy and attachment. Intimacy and attachment have meaning not only for relationships that we traditionally think of as intimate (e.g., between partners and between parents and children) but also for more extended ties. For instance, when relationships are

solid at a community level, individuals feel strong bonds and attachment to places (e.g., neighborhood) and organizations (e.g., voluntary and religious).

Social influence Networks may influence health via several other pathways. One pathway increasingly recognized is social influence. Shared norms around health behaviors (e.g., alcohol and cigarette consumption, sexual practices, and health-care use) might be powerful sources of social influence with direct consequences for the behaviors of network members. For instance, the relationships of adolescent girls with both their male and female friends and with their parents are associated with pregnancy risk. The social influence that extends from the network's values and norms constitutes an important and underappreciated pathway through which networks impact health.

Social engagement and attachment A third and more difficult-to-define pathway by which networks may influence health status is by promoting social participation and social engagement. Participation and engagement result from the enactment of potential ties in real-life activity. Getting together with friends, attending social functions, participating in occupational or social roles, joining in group recreation, attending church are all instances of social engagement. Thus, through opportunities for engagement, social networks define and reinforce meaningful social roles, including parental, familial, occupational, and community roles, which in turn provide a sense of value, belonging, and attachment. Several recent studies suggest that social engagement is critical in maintaining cognitive ability and in reducing mortality.

In addition, network participation provides opportunities for companionship and sociability. Some have argued that these behaviors and attitudes are not the result of the provision of support *per se* but are the consequence of participation in a meaningful social context in and of itself. One reason measures of social integration or connectedness may be such powerful predictors of mortality over long periods of follow-up is that these ties give meaning to individuals' lives by virtue of enabling them to participate fully, to be obligated (in fact, often to be the providers of support), and to feel attached to their community.

Person-to-person contact Another behavioral pathway by which networks influence disease is by restricting or promoting exposure to infectious disease agents. In this regard, the methodological links between epidemiology and networks are striking. What is perhaps most remarkable is that the same network characteristics that can be health-promoting can at the same time be health-damaging if they serve as vectors for the spread of infectious disease. Efforts to link mathematical models applying network approaches to epidemiology are in their infancy and have started to appear over the last 10 years.

The contribution of social network analysis to model of disease transmission is the understanding that in many, if not most cases, disease transmission is not spread randomly throughout a population. Social network analysis is well suited to the development of models in which exposure between individuals is not random but rather is based on geographic location, sociodemographic characteristics (e.g., age, race, and gender), or other important characteristics of the individual such as socioeconomic position, occupation, and sexual orientation. Furthermore, because social network analysis focuses on characteristics of the network rather than on characteristics of the individual, it is ideally suited to the study of the diffusion of transmissible diseases through populations via bridging ties between networks or to uncovering the characteristics of ego-centered networks that promote the spread of disease.

Access to resources and material goods Surprisingly little research has sought to examine differential access to material goods, resources, and services as a mechanism through which social networks might operate. This, in our view, is unfortunate given the work of sociologists showing that social networks operate by regulating an individual's access to life opportunities by virtue of the extent to which networks overlap with other networks. In this way, networks operate to provide access or to restrict opportunities in much the same way as the social status does. Perhaps the most important among studies exploring this tie is Granovetter's classic study of the power of weak ties that, on the one hand, lack intimacy but, on the other hand, facilitate the diffusion of influence and information and provide opportunities for mobility.

Conclusion

This article has reviewed the ways in which social isolation, lack of support, and disengagement may be chronically stressful social experiences and are linked to poor health outcomes. With the development of multilevel framework, we are struck by two issues of profound importance. The first is the upstream question of identifying the conditions that influence the development and structure of social networks. Such questions have been the substantive focus of much of social network research, especially in relationship to urbanization, social stratification,

and culture change. Yet little of this work has been integrated with health issues in a way that might guide us in the development of policies or interventions to improve public health. Recent work relating social cohesion to economic inequality has begun to help us decipher the complex interrelationships between these social experiences. Of particular interest are the ways in which social policies may impact the form and function of social networks. For instance, work/family policies and housing policies may impact network and supportive relationships in important ways, ultimately having a health impact.

The second major issue relates to the downstream question. Many investigators have assumed that networks influence health via social support functions. Our framework makes clear that this is but one pathway linking networks to health outcomes. Furthermore, the work on conflict and stress points out not only that are not all relationships positive in valence but that some of the most powerful impacts on health that social relationships may have are through acts of abuse, violence, and trauma. Fully elucidating these downstream experiences and how they are linked to health and via which biological mechanisms remains a major challenge in the field.

Acknowledgments

This paper is adapted from Berkman, L. F. and Glass, T. (2000), Social integration, social networks, social support and health, In Berkman, L. F. & Kawachi, I. (eds) *Social epidemiology*. New York: Oxford University Press; and Berkman, L. F. (2001), Social integration, social networks, and health. In Smelser, N. J. & Baltes, P. B. (eds), *International encyclopedia of the social and behavioral sciences*. New York: Elsevier.

See Also the Following Article

Social Support.

Further Reading

Bassuk, S., Glass, T. and Berkman, L. F. (1999). Social disengagement and incident cognitive decline in community-dwelling elderly persons. *Annals of Internal Medicine* **131**, 165–173.

Berkman, L. (2001). Social integration, social networks and health. In: Smelser, N. J. & Baltes, P. B. (eds.) *International encyclopedia of the social and behavioral sciences* (vol. 21), pp. 14327–14332. New York: Elsevier.

Berkman, L. F. and Glass, T. (2000). Social integration, social networks, social support and health. In: Berkman, L. F. & Kawachi, I. (eds.) *Social epidemiology*, pp. 137–173. New York: Oxford University Press.

Berkman, L. F. and Syme, S. (1979). Social networks, host resistance, and mortality: a nine-year follow-up of Alameda County residents. *American Journal of Epidemiology* **109**, 186–204.

Bearman, P. and Bruckner, H. (1999). Peer effects on adolescent sexual debut and pregnancy: an analysis of a national survey of adolescent girls. The National Campaign for the Prevention of Teen Pregnancy.

Cohen, S., Underwood, S. and Gottlieb, B. (2000). *Social support measures and intervention*. New York: Oxford University Press.

Durkheim, E. (1897). *Suicide*. New York: Free Press.

Fischer, C. S., Jackson, R. M., Steuve, C. A., et al. (1977). *Networks and places*. New York: Free Press.

Glass, T., Mendes de Leon, C., Marottoli, R., et al. (1999). Population based study of social and productive activities as predictors of survival among elderly Americans. *British Medical Journal* **319**, 478–483.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology* **78**, 1360–1380.

Hall, A. and Wellman, B. (1985). Social networks and social support. In: Cohen, S. & Syme, S. L. (eds.) *Social support and health*, pp. 23–41. Orlando, FL: Academic Press.

Kawachi, I., Colditz, G. A., Ascherio, A., et al. (1996). A prospective study of social networks in relation to total mortality and cardiovascular disease in men in the U.S.A. *Journal of Epidemiology Community Health* **50**, 245–251.

Pennix, B. W., van Tilburg, T., Kriegsman, D. M., et al. (1997). Effects of social support and personal coping resources on mortality in older age: the Longitudinal Aging Study, Amsterdam. *American Journal of Epidemiology* **146**, 510–519.

Seeman, T. (1996). Social ties and health: the benefits of social integration. *Annals of Epidemiology* **6**, 442–451.

Sugisawa, H., Liang, J. and Liu, X. (1994). Social networks, social support and mortality among older peoplein Japan. *Journal of Gerontology* **49**, S3–S13.

Thoits, P. (1995). Stress, coping, and social support processes: where are we? What next? *Journal of Health and Social Behavior* (extra issue), 53–79.

Social Status and Stress

D de Ridder
Utrecht University, Utrecht, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D de Ridder, volume 3, pp 468–473, © 2000, Elsevier Inc.

Social Status, Health, and Stress
Indicators of Social Status
Social Status and Exposure to Stressful Conditions
Social Status and Vulnerability to Stressful Conditions

Glossary

Coping resources	The personal (e.g., self-esteem and sense of control) and social (most notably, social support) resources an individual has at his or her disposal to counter the adverse effects of the experience of distress.
Coping strategies	The cognitive and behavioral attempts an individual may employ when faced with the experience of distress.
Exposure	The actual confrontation with stressful situations (events) of chronic stressful conditions.
Social status	Any attribute indicating the social position of an individual in a group or in the society. The most important attribute is socioeconomic status, comprising income, education, and occupation.
Vulnerability	The lack of personal and social coping resources, which increases the risk of a greater adverse impact of stressful conditions.

Social Status, Health, and Stress

The relationship between social status and health is intriguing because it challenges existing models of how a distal variable such as social status might affect health. Research in the past decades focused on explaining the association between low social status, especially low socioeconomic status (SES), and poor health, assuming that the greater distress of living in deprived social and material conditions is an important factor in the strong and consistent association between health and social status. The extensive literature on SES and health provides some indications why the theoretical framework of stress should be considered an interesting candidate in furthering our understanding of the association between SES and health. First, the linearity of the relationship between SES and health implies better health outcomes as individuals ascend the SES continuum, so that even relatively affluent groups exhibit worse health than their higher SES counterparts. This suggests that factors other than the obvious role of inadequate financial resources or poor living conditions are involved and that SES might also affect psychosocial factors. Identifying these factors may be considered a search for more health-proximal processes so as to bridge the gap between indicators of social status and the associated health outcomes. Second, the analysis of potential measurement artifacts shows that the association between social status and health is robust and is little affected by social selection factors, which would be the case if those who are in poor health were at a higher risk of getting a lower social status. Social selection effects are rare, however, and limited to young children who are unable to attend school as a result of a serious and disabling disease. The absence of social selection effects suggests that the strong association between social status and health is explained by social causation mechanisms – social status affects health instead of the other way around. Particular elements of the stress-coping paradigm offer a framework to study the causal chain between social status at one end of the continuum and health at the other one, particularly those elements that deal with conceptualizing environmental risk for health (exposure to stressors) and individual responses to explain the impact of these stressors (vulnerability to stress).

Indicators of Social Status

Traditionally, most social-epidemiological research has employed the criteria of income, occupation, and education for assessing social status. Less frequently, indicators of material wealth such as car ownership, house ownership, the number of computers in the house, or the number of annual holidays have been used. Interestingly, the sociocultural component of SES expressed in educational attainment appears to have a stronger impact than the more socioeconomic elements of income and occupational status. Recent research indicates that the relationship between income and health varies substantially by level of education. Education improves health, and its effects are larger at low levels of income; education

also reduces the strength of the income–health relationship. The linear gradient relationship between income and health is thus more characteristic of groups with higher levels of education, suggesting that education might improve skills needed to cope effectively with stress. This is shown, for example, in the finding that the mother's educational status is a better predictor of family stress than both family income or the occupational status of the father.

Recently, it has been argued that it is not the absolute level of SES that is important for understanding its association with health and distress but rather inequality from relative standing. There is suggestive evidence that subjective social status (assessed by a ladder with 10 rungs representing where people stand in society in terms of who are the best off, have the most money, have most education, and have the best jobs) is more strongly linked to health in terms of self-rated health, heart rate, and cortisol habituation to repeated stress than the traditional objective measures of SES (education, income, and occupation). The so-called hierarchy stress perspective argues that the gradient-like association between SES and health reflects social ordering rather than material deprivation *per se*. According to this perspective, higher income leads to less stress, less status anxiety, and more perceived control. These underlying mechanisms are psychological and related to relative rather than absolute material deprivation. Social ordering has effects that are no less substantial than absolute deprivation because relative deprivation leads to worse health both directly through physiological pathways (e.g., cortisol levels) and indirectly through maladaptive coping behaviors (e.g., smoking or alcohol consumption). It is noteworthy that in countries with a high degree of income inequality (e.g., the United States) the SES–health gradient is steeper than in economically more egalitarian European countries. Thus, although SES is a reliable and robust indicator of health, SES as such provides few cues for understanding its relationship with stress. In the words of the social epidemiologists Brown and Harris (1978; p. 10), "It is not that the 'demographic type' measures are of no use, it is that they are not enough. What is required is their combination with concepts and measures dealing directly and in detail with the immediate . . . experience of the individual."

Such an approach is found in social role research, advancing the hypothesis that the lower social status of, for example, the nonemployed affects their sense of meaning and purpose in life and has much less to do with the direct material consequences of their income and employment status. Striking findings in this area have been reported 30 years ago on the disadvantaged social role of nonemployed married women, who generally reported poorer health than both employed married women and men – although it is not clear to what extent social changes since the 1970s concerning the role of women make those observations still valid. These studies revealed that women who confront the double role of parent and employee reported lower levels of psychological distress than women who are devoted to the role of homemaker. These findings suggest that regardless the actual level of stressful experiences due to overload, the social status of worker is somehow beneficial for the employed women's health. It should be noted, however, that, although working women may be more protected from distress than nonemployed women, they are still confronted with more distress than employed men.

The results from social role studies are corroborated by findings on particular social groups in our Western community who cannot claim high social status regardless of their income, education, or employment status. Research on the elderly, the unmarried, and nonwhites puts results from social role studies in a somewhat different perspective, suggesting that it is not only the meaning associated with particular roles that accounts for the greater distress of those low in social status but also their ranking order in the social group. Individuals from disadvantaged social groups have been reported to be in poorer health and facing more distress, which may be accounted for by their lower social ranking, although direct tests of this assumption in humans are not available. Animal studies have shown that subordination tends to be associated with a chronically overactive stress response (e.g., higher blood pressure). This pattern is believed to reflect the classic picture of dominance hierarchies as linear pecking orders in which resources are unevenly distributed, inequalities are maintained through aggression and intimidation, and subordinates are subject to the most severe resource limitations, the fewest opportunities for coping, and the greatest physical and social stressors. However, recent studies suggest that the relationship between social rank and patterns of the stress response may be more diverse than has been assumed hitherto. More specifically, it appears that the negative effects of low ranking on stress responses may be attenuated by complex affiliative relationships that may help low-status stressed individuals to cope. It is unknown, however, to what extent these studies on social ranking in animals are directly relevant for humans because humans often belong to multiple hierarchies and tend to value most the one in which they rank highest.

Social Status and Exposure to Stressful Conditions

The common approach to social status and psychological distress suggests that those who are in a disadvantaged social position are confronted with more stressors than their more privileged counterparts, either as a result of exposure to more frequent or more intense negative life events or of living in chronic stressful conditions associated with continuous exposure to daily hassles. Thus, this approach reflects the notion that poorer health associated with lower social status results from social causation, emphasizing the role of structural and material conditions in causing distress. The life event approach to stress dictates that major life events (e.g., divorce or being fired) impose a breach in behavioral routines, cause distress, and therefore require adaptive efforts, eventually exhausting the system. The chronic stress approach takes a somewhat different perspective and states that is the repeated exposure to minor events or daily hassles (e.g., refused services or poor transportation facilities) that ultimately causes wear and tear on the individual.

Life Events

The assumption that people with low social status report more life events than those with higher status is corroborated only weakly by empirical research. Insofar as individuals of lower SES have reported higher numbers or life events, most of these events are directly related to financial problems. There is also a higher report of uncontrollable life events, reflecting to some extent the conditions of poverty and deprivation (e.g., discharge due to becoming obsolete) in individuals with lower social status. Still, even when a higher frequency of life events is reported, this generally fails to account for the higher levels of self-reported distress; no more than 10% of the variance in distress is explained by exposure to life events. Adopting an approach that allows for a contextual evaluation of events shows slightly different results, demonstrating that it is not so much the number of events or even their quality (e.g., uncontrollability) that counts but their potential influence on a number of crucial life areas. Research by social epidemiologists Brown and Harris describing the onset of depression in working-class women clearly demonstrated that these women were confronted with life events that had the potential of generating a chain of other stressful events. The most typical event these women reported involved a teenage pregnancy, which caused these young women to drop out of school and engage in an untimely and unwanted marriage, eventually leaving them divorced, socially isolated, financially deprived, and without a significant chance of entering the labor market or controlling their future in another way – leaving them in a state of hopelessness that is characteristic of depression. Other studies of lower-class women have demonstrated that the lives of these women are fraught with a large number of network events, meaning that they consider the life events experienced by significant others in their social network as if they concerned them directly, which may further explain the higher levels of distress reported by these women. Nevertheless, it is clear that, although the number of major life events appears to be somewhat higher in those occupying low social status, the higher levels of exposure to major life events are not a sufficient explanation to account for the higher levels of distress.

Daily Hassles and Chronic Stressful Conditions

It appears that it is not so much the incidence of particular major life events that is causing high levels of distress in low-social-status individuals but the continuous confrontation with a large number of repetitive small events that mark living in deprived circumstances. The lower an individual is on the SES continuum, the greater the amount of hassles and the greater the time needed to address the basic tasks of living, including shopping in poorly provided facilities, lack of good transportation, poor access to social and health-care facilities, and refused services. Of course, such difficulties in accomplishing daily routines have their echo in social relations. It is not surprising, therefore, that minor conflicts in relationships in work, marriage, and parenting appear to be more common in low-SES individuals. Still, it is hard to imagine that the experience of minor stress is a simple question of individual exposure. Improving the understanding of how chronic stress affects individuals of lower SES may not so much lie in the study of individual experience of these minor stressors but in identifying the environmental features of low-SES neighborhoods themselves, implying that low SES may have an impact on a community level too. Interestingly, research shows that neighborhood-level indicators predict distress above the effects of individual-level indicators. This suggests that the more proximal environment in which people live may provide important information for identifying the environmental features that cause distress. Those in the lower ends of the social-class distribution disproportionally live in deprived neighborhoods, occupy jobs characterized by high demands and low control, and live in social environments in which they are exposed to violence, conflict, abuse, crowding, and noise. All these environmental features may be considered to increase the experience of distress, although it

remains to be answered how these unhealthy environments get under an individual's skin (Taylor & Repetti, 1997). Adopting a perspective that conceptualizes social status as a proxy to understanding the stressful features of particular environments raises the issue of how situational demands should be distinguished from subjective appraisals of these objective demands. This issue constitutes a classic problem in the cognitive approach to stress formulated by Richard Lazarus and deals with the important question of how demands in the environment are appraised and handled by the individual.

Social Status and Vulnerability to Stressful Conditions

No matter how unhealthy and environment is, not all individuals sharing the same environment are affected by that environment in a similar way, suggesting that factors concerning individual susceptibility to these environmental threats are at work. The approach highlighting a differential vulnerability to stressful conditions may improve our understanding of why individuals of lower social status experience more distress. Vulnerability is reflected in the availability of resources that may be helpful in countering the adverse effects of exposure to stress, including the repertoire of coping strategies, social support, and beliefs about control and competence. Because exposure to environmental threats fails to offer a sufficient explanation for the experience of greater distress in low-SES individuals, it has been hypothesized that the availability of coping resources varies with social status. The question is whether and why lower SES is associated with lower availability of these and other coping resources.

Coping

The employment of adequate cognitive and behavioral coping efforts constitutes one of the most important resources for buffering the adverse effects of psychological distress. Individuals who find constructive ways of coping with stress, such as taking direct action or finding meaning in their experience, are better able to withstand the negative effects of stressful circumstances than individuals who believe their coping resources to be minimal or who believe that the situation is beyond their control and rely on avoidance coping strategies, which in the long run often turn out to compromise their health. Generally, people of low social status are more inclined to adopt avoiding and emotion-focused strategies when they are confronted with stress than are their more advantaged peers. Education level and family income directly relate to the frequency of problem-focused coping for dealing with financial strain and other hassles, with those of low social status reporting less problem-focused coping. Why low SES promotes generally ineffective coping strategies is unknown, although it has been suggested that the acquisition of coping strategies may to some extent be shaped by the continuous confrontation with few possibilities to control their environment. Although research that directly links social status and coping is rare, it has been shown that individuals of low social status are more sensitive to stress and think of it as unpleasant, disruptive and beyond their control. In contrast, individuals with higher social status appear to be less bothered by distressing events and more often regard it as a challenge to do something about it. It thus seems that the appraisal of potentially upsetting situations as uncontrollable and overwhelming may activate dysfunctional coping patterns. More generally, it seems that low-SES environments reduce individuals' reserve capacity to manage stress, thereby increasing negative emotions and cognitions with subsequent effects for dealing with these stressful conditions themselves.

Social Support

The adverse effect of chronic exposure to stressful environments may be partially offset if an individual has at least one supportive person in his or her neighborhood. Living in low-SES communities may compromise the effective use of social contacts, however. Although the social network of low-SES individuals generally is not smaller than the network of those with better education and higher incomes, the networks of low-SES individuals tend to be more homogeneous, with a large proportion of kin, which may affect the type of support provided and its perceived helpfulness. If an individual is not willing to share particular problems with family members, the possibilities for social support decrease dramatically. The larger proportion of kin in the social networks of low-social-status individuals also has consequences for mutual dependency. Asking for help on one occasion implies some kind of obligation to provide support when a family member asks for it on another occasion. The larger proportion of family in the network also bears costs because it may promote the experience of stress when a family member experiences distress. In addition, the rare prospective research in this area demonstrates that the experience of distress may affect the availability of social support. Studies of economic stressors such as financial hardship or unemployment confirm a pattern in which economic stressors lead to increased marital stress and decreased social support.

Social Distribution of Coping Resources

It is not clear why low-SES individuals are at a disadvantage with regard to the availability of coping resources. Vulnerability is thought to be partly genetically determined and partly acquired during childhood and adolescence by accumulating negative experiences, which may be difficult to compensate for. Lack of financial resources and poor material conditions are, of course, more common in low-SES individuals, but these factors do not appear to be the most important determinants of vulnerability. No evidence is available that limited access to coping resources is a direct result of deficient life conditions. Low self-esteem and low personal control are acquired during repetitive and accumulating negative events causing a negative spiral in which an individual learns to be helpless and hopeless. Low chances of employment or low educational attainment do not by themselves cause vulnerability to distress, but this does not mean that vulnerability does not operate at a social level. Particular resources and coping strategies may be distributed differently among the social classes and also particular beliefs about stress and health may be socially shaped and transmitted. The employment of ineffective coping strategies seems to depend on strong beliefs about what constitutes stress and what can be done about it. Beliefs reflect, at least in part, socially shared assumptions about stress, coping, and health and may be considered cultural models of the individual and his or her social group. A significant example concerns cigarette smoking as a way of coping. Although smoking is widely recognized as a health-threatening habit in both low-SES and high-SES individuals, in low-social-status individuals smoking has often turned into one of their few available coping options and creates a feeling of rest and control when faced with distress. Further research on the social distribution of coping resources is certainly required.

See Also the Following Articles

Aging and Psychological Stress; Cultural Factors in Stress; Economic Factors and Stress; Education Levels and Stress; Marital Status and Health Problems; Minorities and Stress; Smoking and Stress; Social Capital.

Further Reading

Abbott, D. H., Keverne, E. B., Bercovich, G. B., et al. (2003). Are subordinates always stressed?: a comparative analysis of rank differences in cortisol levels among primates. *Hormones and Behavior* **43**, 67–82.

Adler, N. E., Boyce, T., Chesney, M. A., et al. (1994). Socioeconomic status and health: the challenge of the gradient. *American Psychologist* **49**, 15–25.

Adler, N. E., Epel, E. S., Castellazzo, G., et al. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. *Health Psychology* **19**, 586–592.

Antonovsky, A. (1967). Social class, life expectancy and overall mortality. *Millbank Quarterly* **45**, 31–73.

Brown, G. W. and Harris, T. (1978). *Social origins of depression: a study of psychiatric disorder in women.* London: Tavistock.

Chen, E., Matthews, K. A. and Boyce, W. T. (2002). Socioeconomic differences in children's health: how and why do these relationships change with age? *Psychological Bulletin* **128**, 295–329.

Cohen, S., Kaplan, G. A. and Salonen, J. T. (1999). The role of psychological characteristics in the relation between socioeconomic status and perceived health. *Journal of Applied Social Psychology* **29**, 445–468.

Feldman, P. J. and Steptoe, A. (2004). How neighborhoods and physical functioning are related: the roles of neighborhood socioeconomic status, perceived neighborhood strain, and individual health risk factors. *Annals of Behavioral Medicine* **27**, 91–99.

Gallo, L. C. and Matthews, K. A. (2003). Understanding the association between socioeconomic status and health: do negative emotions play a role? *Psychological Bulletin* **129**, 10–51.

Sapolsky, R. M. (2004). Social status and health in humans and other animals. *Annual Review of Anthropology* **33**, 393–418.

Taylor, S. E. and Repetti, R. L. (1997). Health psychology: what is an unhealthy environment and how does it get under the skin? *Annual Review of Psychology* **48**, 411–447.

Wilkinson, R. G. (1999). Health, hierarchy, and social anxiety. In: Adler, N. E., Marmot, M., McEwen, B. & Stewart, J. (eds.) *Special issue on socioeconomic status and health in industrial nations: social, psychological, and biological pathways, Annals of the New York Academy of Sciences*, 48–63.

Social Stress, Animal Models of

J Bugajski and A Gądek-Michalska
Polish Academy of Sciences, Krakow, Poland
A J Bugajski
Jagiellonian University, Krakow, Poland

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Bugajski, volume 3, pp 474–478, © 2000, Elsevier Inc.

Social Stress Models
Limbic Stress Processive Pathways
Adaptation to Stress
Transmitter Systems Adaptation
CRH Systems Adaptation
Vasopressin Systems Adaptation

Glossary

Adaptation	The situation in which repeated exposure to the same stimulus elicits diminishing responses from the body.
Corticotropin-releasing hormone (CRH)	Neurohormone secreted by the hypothalamus whose functions are related to stress, a primary activator of stress hormones.
Limbic-hypothalamic-pituitary-adrenal (LHPA) axis	The pathway regulating the secretion of stress hormones by brain limbic structures, hypothalamus, pituitary, and adrenal gland and involving release of corticotropin-releasing hormone (CRH) from the hypothalamus, adrenocorticotropic hormone (ACTH) from the pituitary gland, and corticosteroid hormones from the adrenal gland.
Neurotransmitters	Chemical messengers released from different kind of nerve endings involved in the transmission of a nerve impulse.
Social stress	A highly complex set of neuroendocrine changes in response to the immediate social environment of the animal.
Vasopressin (AVP)	Neurohormone secreted by the hypothalamus, a major regulatory peptide, with CRH, for the secretion and synthesis of ACTH; initiates the hypothalamic-pituitary-adrenal (HPA) response to stressors.

Conventional models of physical stressors in laboratory animals show little validity with chronic human psychopathologies. Models of social stress based on the naturalistic events of interactions between conspecifics present a realistic model to include chronic psychosocial stress in animals. Repeated or chronic psychosocial stress induce adaptation of various neurotransmitter and neuropeptide regulatory systems involved in the hypothalamus-pituitary-adrenal (HPA) axis response. Chronic stress impairs the cholinergic and adrenergic system-induced HPA response. The corticotropin-releasing hormone (CRH)- and vasopressin (AVP)-induced activation of the HPA axis follows distinct time courses. Complex transcription factor interactions and mechanisms accompany the activation of these neuropeptides during psychosocial stress.

Social Stress Models

The stress response is usually presented as adaptive physiological and psychological processes activated in animals and humans by various kinds of stressors. It is regarded as an alarm system that is initiated whenever the dynamic equilibrium of the living organism, called homeostasis, is threatened by physical and psychological events that are known as stressors. Current understanding of physiological mechanisms of the central stress response system is based on both clinical data and animal reactions to various physical stressors, such as footshock, forced swim, or immobility, which may elicit behavioral and psychological responses different from those evoked by psychosocial stressors. In socially organized mammals, however, the predominant stressors are not physical events but arise from the immediate social environment of the animal. The majority of stress stimuli in humans is of a psychosocial nature and contributes to the development and expression of disease. In order to study mechanisms involved in the etiology of human affective disorders, various animal models of social stress are used. Apart from genetic factors that predispose for psychopathologies, environmental stress plays an important role in the etiology of these mental diseases.

Common rodent and nonhuman primate models of social stress used in the laboratory with a focus on social hierarchy models contribute in complementary ways to the understanding of social stress. Animal models of social stress involve single, intermittent, or chronic exposure of a subject animal to a conspecific. The most frequently used models for rodents are the social defeat and the social colony models. In the social defeat model, a male intruder is introduced into the territory of another male, the resident

conspecific, which attacks and defeats the intruder. The desired outcome of the social conflict is ensured since the resident usually has higher body weight, is familiar with fighting, and frequently belongs to a more aggressive strain than the intruder.

In a social colony, male and female rats are living together in seminatural conditions and a social structure develops, with a single male being the dominant rat and the others subordinates. Social defeat stress induces major functional and structural changes in subregions of the hippocampal formation. Defeated or subordinate animals have an increased activation of HPA axis, demonstrated by an increase in blood adrenocorticotropic hormone (ACTH) and corticosterone levels. Consequently, animal models that involve natural social context seem to be more appropriate because they represent situations that individuals may encounter in their everyday lives.

Crowding typically induces a social stress reaction with prominent psychosocial components, avoiding the physical components present in most typically used stressors and mimicking emotional state alterations. Randomly selected animals are placed together in housing cages such that each has less than a standard amount of space. In crowded rats, no agonistic interactions leading to dominance or subordination are encountered. The prototypical marker of the presence and magnitude of social stress is the activity of the HPA axis. Many responses to social stress appear to be quite consistent across various animal models, with only a few exceptions that may depend on the paradigm or species studied. Psychological stress is a good animal model for investigating the various behavioral and neurohormonal changes related to stressful circumstances.

Psychological stressors are associated with the activation of the sympathetic and adrenomedullary system and the HPA axis. These systems respond to psychological stress with increased output of noradrenaline from the locus coeruleus (LC) and noradrenaline and adrenaline from the adrenal medulla. The hypothalamus produces CRH, which activates the anterior pituitary gland to secrete ACTH, which in turn stimulates the adrenal cortex to secrete corticosteroids, primarily cortisol in humans and corticosterone in rodents. Negative emotional responses disturb the regulation of the HPA system. Relatively pronounced HPA activation is common in depression, with episodes of cortisol secretion being more frequent and persistent among the depressed than among other psychiatric patients and normal subjects. Chronic stress may induce a state of hyporesponsiveness of the HPA axis whereby cortisol secretion is attenuated.

Limbic Stress Processive Pathways

Psychological and neurogenic stressors requiring higher-order sensory processing of signals prior to initiation of a stress response recruit the limbic stress processive pathway. These processive stressors constitute stimuli that become stressful only by comparison with previous experience. Different types of emotional experience may involve different pathways and arise from activity initiated as a result of thoughts and memories in the cortex or as a result of activity initiated in the hypothalamus. The components of the limbic-sensitive pathways include prefrontal cortex, septum, extended amygdala, hippocampus, and hypothalamic nuclei. In contrast, the limbic-insensitive pathway is recruited by stressors that represent an immediate threat to homeostasis and does not require limbic structures. The amygdala and the hippocampus normally inhibit hypothalamic activity, but stress blocks this circuitry and thus leads to stimulation of hypothalamic releasing hormones such as the stress peptide. Both the hippocampus and the medial prefrontal cortex (mPFC) play an important role in the negative feedback regulation of HPA activity during physiological and behavioral stress. Chronic restraint and psychosocial stress can induce atrophy of hippocampal apical dendrites of CA3 pyramidal neurons, depletion of synaptic vesicles in mossy fiber terminals in CA3, and impairment of hippocampal-dependent learning tasks. Chronic elevations in corticosterone also induce a suppression in synaptic plasticity in the CA1 hippocampal field.

The amygdala is also a major extrahypothalamic source of CRH-containing neurons with high expression levels for the CRH-1 and CRH-2 receptors. Psychological stressors can activate the amygdala CRH system without evident activation of the hypothalamic CRH system; the amygdala CRH system is much more sensitive to psychological stressors than the hypothalamic paraventricular nucleus (PVN). The amygdala, in concert with other extrahypothalamic limbic regions, manifests the behavioral responses to stress, while the hypothalamic CRH system is important for modulating neuroendocrine responses to stress. The hippocampus is an important target of circulating adrenal cortical hormones that act on their receptors and are abundant in this structure. Stress affects glucocorticoid receptor and mineralocorticoid receptor gene expression in the hippocampus in a stressor- and gender-specific manner. Chronic exposure to high corticosteroid levels, during prolonged stress, can have a detrimental effect on hippocampal integrity and function and its excitation and plasticity.

Adaptation to Stress

Adaptation of a response to given challenge is one of the key features of the normal ability of an organism to cope with the external world. Social factors, as unavoidable real life events, are powerful modulators of the stress response. Adaptation to stressful challenges involves activation of neural, neuroendocrine, and neuroendocrine-immune mechanisms. The ability of the body to increase or decrease vital functions to a new state is termed allostasis, or stability through change. Allostasis actively maintains homeostasis and promotes adaptation. When the stress response is inadequate or excessive and the cost of restitution of homeostasis is too high, the resulting condition has been termed by McEven allostatic load.

Physiological stress responses are accepted as beneficial in the short term but can be detrimental if overactivated or prolonged. Stress modulators regulate the extent, duration, and long-term impact of acute stress responses. Stress hormones are crucial mediators of both adaptive and maladaptive responses. Chronic stress induces changes in neuroendocrine and neuronal systems at several levels: in morphology of brain nuclei and neurons and in neurotransmitter and second messenger systems including single molecules. This also takes place in limbic brain structures that are under the influence of stress-activated systems. Within the limbic-hypothalamic-pituitary-adrenal (LHPA) system, chronic stress impairs the feedback-control mechanisms associated with decreased gene expression of corticosteroid receptors and structural alterations in the hippocampal formation. Prolonged stress reduces the inhibitory feedback exerted by the hippocampus over the hypothalamic CRH-producing PVN and consequently results in hypersecretion. Sustained stress leads to adrenal hyperactivity. LHPA axis alterations develop in all stressed animals independently of their position in the social hierarchy, whether dominants, subordinates, residents, or intruders, indicating that the HPA axis is more sensitive to the stressful nature of the situation and less modulated by the individual differences in the appraisal of the situation. Dysfunction of systems regulating CRH activity in the central nervous system (CNS) and LHPA axis dysregulation often contributes to the etiology of several psychiatric disorders, such as anxiety, panic disorder, posttraumatic stress disorder, and depression.

Challenging signals from the environment are communicated to the final cellular level of adaptation to stress to turn cascades of gene expression. Genes turned on by stress transcribe the information in their DNA code into messenger RNA (mRNA), which then is translated into proteins making up the structure and adaptive functions of the brain and body. Proteins regulate the structural and energy dynamics of the cell, as well as the neurotransmitters, hormones, and signals of the psychoneuroendocrine system in mind–body medicine. A significant number of the genes in the human genome can be spliced together in different ways to generate alternative sets of proteins that modulate physiological adaptation in response to challenge and stress from purely physical to psychosocial levels.

Transmitter Systems Adaptation

Cholinergic System

Acute stress activates not only the HPA axis but also the septohippocampal cholinergic system in male rats. In the brain, acetylcholine (ACh) is one of the neurotransmitters clearly involved in the stimulation of hypothalamic release of CRH. Acute stress of immobilization increases the synthesis, binding, and release of ACh in the hippocampus and limbic structures. Chronic stress induces reduction of choline uptake and an increase in the number of muscarinic binding sites in the septohippocampal cholinergic system, which is involved in the pituitary-adrenal response to stressful stimuli. The social stress of crowding (21 rats placed in a cage designed for 7) for 3, 7, and 14 days considerably reduces the secretion of corticosterone induced by ICV stimulation of cholinergic muscarinic receptors. The reduced responsiveness of the HPA axis may be caused by a homologous desensitization and downregulation of central muscarinic receptors by an increased secretion of ACh during crowding stress. Homologous desensitization of muscarinic receptor may be regulated by β-adrenergic receptor-like kinase.

The function of the cholinergic system depends on the ACh gene, which codes for ACh and the acetylcholinesterase gene, which codes for the enzyme that hydrolyzes ACh and terminates the function of ACh excitatory neurotransmission in the CNS. Psychosocial stress can alter the cholinergic gene expression or splicing to generate alternative proteins that are involved in a variety of normal mind–body functions, as well as pathologies. Such stress-induced alternative gene splicing may constitute a major pathway in the modulation of gene expression by creative psychological and social processes.

Monoaminergic System under Stress

An important component of the stress response in the brain is the activation of monoaminergic systems.

The LHPA axis interacts with a variety of monoaminergic systems, including brain stem noradrenergic and serotonergic systems. Endocrine and central brain stem monoaminergic responses to stress appear to be precisely coordinated to facilitate adaptive physiological responses. Stress-induced stimulation of the noradrenergic/adrenergic, the serotonergic, and the dopaminergic system enables the brain to flexibly respond to challenges by reorganizing its neuronal networks. Acute stress stimulates noradrenaline (NA) release from terminal neurons in the LC, leading to depletion of NA, but repeated or prolonged stress restitutes NA levels in the brain.

Stress raises activity of phenylethanolamine-N-methyltransferase (PNMT), a specific enzyme that converts NA into adrenaline and accelerates turnover of brain NA and adrenaline.

In addition, stress-induced increased expression of the tyrosine hydroxylase gene may elevate NA concentrations. The release of catecholamines in the response to stressors is followed by an increase in the expression of genes that encode catecholamine-synthesizing enzymes by transcriptional mechanisms in the LC and the adrenal medulla. The persistence of transcriptional activation depends on the duration and severity of the stress. In response to an acute stressor, catecholamines stimulate some limbic regions, such as the hippocampus, amygdala, and hypothalamus, whereas the higher cognitive centers in the prefrontal cortex are inactivated. During stress, high concentrations of NA and/or adrenaline cause a decrease in a number of adrenergic receptors, termed downregulation, and low agonist concentrations upregulate adrenergic receptors. Adrenoreceptor (AR) downregulation may result from reduced expression of the receptor gene or proteolytic degradation of receptor molecules. A reduced number of functional receptors results in desensitization of the respective receptor system and impairs its biological response to an agonist. Receptor molecules may be desensitized by phosphorylation through protein kinases, rapid uncoupling from its effector unit, sequestration from the cellular membrane, and internalization into the cytoplasm and intracellular vesicles. Chronic psychosocial stress reduces the number of adrenergic receptors in brain regions that regulate emotional behavior depending on the duration of the stress period and brain regions. In the LC, α_2-AR binding is already reduced 2 days after the onset of the stress and remains at a low level. Diminished autoreceptor activity impairs feedback inhibition of NA release and, consequently, higher levels of the neurotransmitter in projecting regions of the LC. During chronic psychosocial stress, α_2-AR expression alters from initial downregulation to later upregulation, inducing first high and then low levels of NA.

Stress-induced downregulation of brain β-ARs is regarded as an important mechanism of adaptation to stress. Repeated or chronic stress induces high levels of endogenous agonist NA, and adrenaline causes rapid β-AR internalization into the cells and their functional desensitization. This is considered one of the biochemical mechanisms in receptor adaptation performed in order to prevent dangerous effects of the persisting stress-induced high levels of catecholamines. The psychosocial stress of crowding for 3–7 days considerably reduces the secretion of corticosterone elicited by ICV β-adrenergic receptor agonist and evokes progressively lesser diminution of that response after 14 and 21 days of crowding. In brain cortex α_2-ARs are temporarily downregulated after 2 days of stress but upregulated after 4 weeks. Stress rapidly desensitizes the β-AR system via receptor sequestration and internalization, but some of these changes are only transient since internalized β-ARs can reinsert into the plasma membrane and restore normal receptor numbers. Social crowding stress only moderately diminishes the secretion of corticosterone evoked by ICV α_1-adrenergic agonist. Adrenergic receptors, like the cholinergic muscarinic receptor, are linked to G-protein and are regulated by phosphorylation, which regulates desensitization of receptors to their agonists. Phosphorylation of the β-adrenergic receptor by cAMP-dependent kinase and protein kinase C (PKC) during stress is intimately involved in the desensitization of the receptor. The brain, and specifically the hippocampus, have large neuronal plasticity for a successful adaptation to a stressful situation that occurred previously. Changes that occur in the hippocampus are accompanied by alterations in the other brain regions functionally related to the hippocampus. Reacting and changing neurotransmitter systems in brain areas interact in a coordinated fashion aimed at maintaining stability through change in physiological coping mechanisms during stress.

CRH Systems Adaptation

The most well-known hypothalamic peptide whose functions are related to stress is the CRH, which may be regarded as a primary activator of the LHPA system. Produced by many neurons throughout the brain, CRH functions as neurotransmitter. CRH and noradrenergic neurons of the central nervous stress system innervate and stimulate each other in a positive feedback mechanism. CRH given ICV in rats activates the HPA axis and the sympathetic nervous

system. Stress activates hypothalamic as well as extrahypothalamic CRH neuronal pathways. In response to a stressful stimulus, CRH is released into the hypophysial-portal blood from neurons in the PVN of the hypothalamus, the site of convergence for stress signals in the brain. Hypothalamic CRH acts as a major mediator of the mammalian stress response by stimulating pituitary pro-opiomelanocortin (POMC) gene expression and ACTH secretion that, in turn, stimulates release of glucocorticoid from the adrenal gland. CRH stimulates POMC gene transcription in pituitary corticotrops by several transcription factors. The transcriptional activation of the pituitary POMC gene by CRH is associated with inhibition of nuclear transcription factor NF-κB DNA-binding activity and with activation of the activator protein-1 (AP-1) factor. CRH activates target pituitary cells and neurons via distinct genes encoding membrane CRH-R1 and CRH-R2 receptors. Stress causes marked increases in the expression and cell content of CRH-R1 receptors, which predominantly mediate the CRH-induced hormonal response to stress. CRH released within the PVN increases cAMP levels, and sequentially facilitates CRH transcription, via cAMP response-element binding protein (CREB) phosphorylation and later limits the transcription via induction of the inducible cAMP early repressor (ICER). Fine modulation of CRH expression depends on a complex balance between activation by phospho-CREB and repression by ICER late during stress. Induction of ICER could serve as a protective mechanism to limit transcription of hypothalamic CRH and avoid pathological consequences of excessive expression of neuropeptide in response to stress.

Precise regulation of circulating glucocorticoids is critical for preserving homeostasis under stress conditions. Excessive CRH production may lead to HPA axis dysfunction and psychiatric disturbances. The inhibitory isoforms of CRE modulator induced during stress contribute to the decline in CRH gene transcription during persistent stimulation.

Vasopressin Systems Adaptation

In the initiation and modulation of the HPA axis response to acute and psychosocial stress both CRH and AVP are involved. These two major regulatory peptides for the secretion and/or synthesis of ACTH are colocalized in the parvocellular neurons in the PVN in both humans and rodents. These peptides, colocalized in the same secretory granules at the nerve endings of the median eminence and coreleased into the portal circulation, exert synergistic effects on ACTH secretion from the anterior pituitary. Apart from presence in the hypothalamus, CRH and AVP are localized in several other limbic and brain structures. A variety of acute and chronic stressors alters their concentrations in the hypothalamus and several other brain regions.

In pituitary corticotrops, the vasopressin V1b receptor isoform is specifically expressed and plays a crucial role in regulating HPA axis activity during stress and in resting conditions. Under stress conditions, both the hypothalamic AVP/V1bR and the CRH/CRH R1 system are required to maintain normal HPA axis regulation. Acute stress results in a rapid increase of CRH heterogeneous nuclear RNA (hnRNA) at the time of maximal ACTH secretory response. The stress-induced AVP hnRNA response in the parvocellular neurons is slower in onset than that of primary CRH transcript, with a maximum at 2 h after challenge. Differences in timing of CRH and AVP hnRNA responses indicate distinct mechanisms underlying stress-induced activation of these genes. Physical or psychological stressors activate the CRH and AVP in the parvocellular neurons of the PVN, the major regulators of ACTH secretion. The AVP gene is upregulated by cAMP. The promoter of the AVP gene contains cAMP response element (CRE), and the promoter activates of transfected AVP gene are upregulated with cAMP. The dual CRH/AVP regulatory system may benefit the organism in coping with various stressful events. The release of CRH and AVP peptides is regulated differently depending on the nature and the duration of the imposed stress. Prolonged stressors induce a gradual shift from a CRH-predominant stress response to an AVP-predominant response after chronic stress. These adaptational changes could partly depend on glucocorticoid receptor changes induced by repeated stress in the PVN and the hippocampus and the greater responsivity of AVP than CRH to chronic stimulatory inputs. CRH and AVP regulate MR and GR in the hippocampus and anterior pituitary, structures with the highest levels of MR and GR mRNA and receptor protein in the CNS. These GRs have an important role in the regulation of the LHPA axis and stress adaptation and the pathophysiology of stress-related disorders.

AVP and CRH bind to specific pituitary corticotrop membrane receptors that activate different second messengers. The rat pituitary AVP receptor is coupled to phospholipase C, and the CRH receptor to an adenylate cyclase. Repeated stress induces rapid adaptation of the CRH mRNA response and desensitization of the pituitary ACTH response, which may be caused by negative feedback mechanisms, as well as depletion of the releasable ACTH pool or downregulation of pituitary CRH receptors. On the other

hand, chronic stress or repeated exposure to identical short-term stressors may induce an increased pituitary responsiveness due to sensitization of pituitary corticotrops to CRH. The social stress of crowding for 3 days facilitates the ACTH response to CRH and corticosterone response remains unaffected, indicating that the HPA system is fully sensitive or even hyperactive. This may reflect predominantly moderate upregulation of pituitary CRH system. Crowding stress considerably reduces the AVP-induced ACTH and corticosterone response. This results from desensitization and/or decrease in the number of AVP receptors in the anterior pituitary corticotrops or from desensitization of the AVP-stimulated intracellular inositol triphosphate and the PKC signaling pathway. AVP plays a primary role in the regulation of the pituitary adrenal axis during adaptation to stress. Chronic stress increases the expression of AVP in parvicellular neurons of the PVN, and secretion of AVP but not CRH into the pituitary portal circulation. Changes in density and sensitivity of pituitary AVP receptors may play a major role in adaptation of the HPA axis to crowding stress, as with adaptation to chronic psychosocial stress.

See Also the Following Articles

Animal Models (Nonprimate) for Human Stress; Crowding Stress; Psychosocial Factors and Stress; Sex-Specific Effects of Early Social Stress in Mammals: a Study in Guinea Pigs.

Further Reading

Barr, C. S., Newman, T. K., Shannon, C., et al. (2004). Rearing condition and rh5-HTTLPR interact to influence limbic-hypothalamic-pituitary-adrenal axis response to stress in infant macaques. *Biological Psychiatry* **55**, 733–738.

Bartolomucci, A., Palanza, P., Sacerdote, P., et al. (2005). Social factors and individual vulnerability to chronic stress exposure. *Neuroscience and Biobehavioral Reviews* **29**, 67–81.

Blanchard, R. J., McKittrick, C. R. and Blanchard, D. C. (2001). Animal models of social stress: effects on behavior and brain neurochemical systems. *Physiology and Behavior* **73**, 261–271.

Bugajski, J. (2000). Social stress. In: Fink, G. (ed.) *Encyclopedia of Stress* (Vol. 3), pp. 474–478. San Diego, CA: Academic Press.

Bugajski, J., Borycz, J., Głód, R., et al. (1995). Crowding stress impairs the pituitary-adrenocortical responsiveness to the vasopressin but not corticotropin-releasing hormone stimulation. *Brain Research* **681**, 223–228.

Buwalda, B., Kole, M. H. P., Veenema, A. H., et al. (2005). Long-term effects of social stress on brain and behavior: a focus on hippocampal functioning. *Neuroscience and Biobehavioral Reviews* **29**, 83–97.

Flugge, G., van Kampen, M. and Mijnster, M. J. (2003). Perturbations in brain monoamine systems during stress. *Cell Tissue Research* **315**, 1–14.

Fuchs, E. and Flugge, G. (2003). Chronic social stress: effects on limbic brain structures. *Physiology and Behavior* **79**, 417–427.

McEwen, B. S. (1998). Stress, adaptation, and disease. Allostasis and allostatic load. *Annals of the New York Academy of Sciences* **840**, 33–44.

McEwen, B. S. (2000). The neurobiology of stress: from serendipity to clinical relevance. *Brain Research* **886**, 172–189.

Rossi, E. L. (2004). Stress-induced alternative gene splicing in mind-body medicine. *Advances in Mind-Body Medicine* **20**, 12–19.

Sabban, E. L. and Kvetnansky, R. (2001). Stress-triggered activation of gene expression in catecholaminergic systems: dynamics of transcriptional events. *Trends in Neuroscience* **24**, 91–98.

Tamashiro, K. L. K., Nguyen, M. M. N. and Saki, R. R. (2005). Social stress: from rodents to primates. *Frontiers in Neuroendocrinology* **26**, 27–40.

Tanoue, A., Ito, S., Honda, K., et al. (2004). The vasopressin V1b receptor critically regulates hypothalamic-pituitary-adrenal axis activity under both stress and resting conditions. *The Journal of Clinical Investigation* **113**, 302–309.

Watanabe, Y., McKittrick, C. R., Blanchard, D. C., et al. (1995). Effects of chronic social stress on tyrosine hydroxylase mRNA and protein levels. *Molecular Brain Research* **32**, 176–180.

Social Support

T C Antonucci
University of Michigan, Ann Arbor, MI, USA
J E Lansford
Duke University, Durham, NC, USA
K J Ajrouch
Eastern Michigan University, Ypsilanti, MI, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by T C Antonucci, J E Lansford and K J Ajrouch, volume 3, pp 479–482, © 2000, Elsevier Inc.

Theoretical Perspectives
Social Support as a Mechanism for Coping with Stress
Social Support as a Source of Stress
Summary and Conclusion

Glossary

Anticipated support	Support perceived to be available if the need for it arises in the future.
Appraisal support	Feedback relevant to self-evaluation.
Emotional support	The provision of empathy, reassurance, liking, and respect.
Informational support	Help with problem solving and advice.
Instrumental support	Aid with services and the tasks of daily living.
Social network	The group of individuals with whom an individual may exchange support.

This article examines the concept of social support and its relationship to stress. A brief review of several illustrative theoretical perspectives of social support is presented first. Next, an outline and overview of the research focusing on social support as a coping mechanism is provided, followed by a discussion of social support as a source of stress and a brief summary and conclusion.

Theoretical Perspectives

Social support has received a great deal of attention in the last several decades. There are some points that are now relatively well accepted: social support tends to be cumulative over the life course; it is often multigenerational; it is best when an individual both provides and receives support, although not necessarily in the context of the same relationship; and it can include exchanges between family and friends and between formal and informal providers.

Research thus far has offered interesting and sometimes quite useful evidence for how social relations affect the individual's socialization, development, coping, and general well-being. The concept of social support has been attractive to researchers in a wide variety of disciplines, including epidemiology, sociology, psychology, social work, and health care. The related theoretical perspectives or frameworks offered by each of these disciplines are too numerous to mention here; instead, illustrative examples are provided.

Epidemiologists offered some of the earliest theorizing about the role of social support and health. Researchers in this field suggested that social relations help protect the individual from disease. This perspective was initially considered quite revolutionary because these researchers claimed that physical as well as mental health could be influenced by social relations. More recently, other researchers have found empirical support for this perspective.

Kahn and Antonucci extended this notion by suggesting that individuals are surrounded by convoys of family and friends. The resulting interaction and socialization processes teach them the roles and responsibilities of childhood, adulthood, and old age. This convoy of relations is assumed to have a positive effect, helping the individual successfully meet the challenges of life. However, it is recognized that convoys may not always be positive and sometimes can expose individuals to stress rather than serve as protection. This theoretical framework suggests that social support is best understood within a life course perspective and that social interactions are cumulative and developmental.

These theoretical frameworks were general and broad; other theories have taken a more practical approach to understanding social support. Cantor's hierarchical theory of social support suggests that a natural hierarchy of people exists from whom an individual prefers to receive support. People prefer to receive support from a spouse and children first and then turn to friends and neighbors if the former are unavailable. Only if all of these informal sources of support are unavailable are people likely to seek support from a formal provider (e.g., government and community institutions). A somewhat complementary theory was offered by Litwak, who suggested that people turn to those who are most able to provide the specific goods or services they need rather than to those to whom they feel closest. Both of these theoretical perspectives have a practical task-oriented focus.

Some theories have been more psychologically focused. For example, Pearlin argued that others step in when needed and help the individual achieve his or

her goal. This provides the individual with a feeling of ability or personal achievement, which not only aids in coping with a specific event but also provides the individual with a more generalizeable sense of mastery. This theoretical perspective is similar to that offered by Antonucci and Jackson, who suggested that there is a life-span developmental benefit to the constant repetition of supportive interactions. Receiving support over time leads to viewing oneself as worthy, cared for, and loved. Under optimal conditions, this exchange provides the individual with a sense of personal efficacy that is critical as the individual faces the challenges or crises of life.

The types of theoretical perspectives that have guided research in this area can be summarized as follows: (1) general broad-based theories suggesting that social relations affect an individual's ability to resist stress and disease; (2) more practical, task- or person-specific theoretical frameworks that provide guidelines for determining who is the logical choice for helping a person cope with a specific problem, task, or crisis; and finally (3) more global, developmental perspectives that are designed to explain how an individual develops the ability to cope with stress over time. We turn now to a consideration of the evidence that examines how social support influences the individual's experience of stress.

Social Support as a Mechanism for Coping with Stress

It is well established that social relationships influence individuals' psychological well-being by providing love, intimacy, reassurance of worth, tangible assistance, and guidance. Across the life span, the lack of high-quality relationships is associated with negative physical and psychological consequences such as anxiety, depression, loneliness, and poor health. Studies of social support and well-being generally reflect two lines of research. First, support is thought to have direct effects on health maintenance and recovery, enhancing well-being regardless of the level of stress experienced. Second, support has been conceptualized as having indirect effects on well-being by acting as a buffer against stress. According to this second approach, during stressful events social support mediates the association between stress and adjustment so individuals who have high-quality support are not as negatively affected by stress as those who have less adequate support. Research provides evidence for both models. In this section, however, we focus on conceptualizing support as a buffer against stress.

Empirical studies provide evidence that social support can reduce the negative impact of stress associated with physical illness, unwanted or unexpected life events, and developmental transitions. For example, supportive relationships with family members and friends can help children cope with stress associated with school transitions, young adults adjust to the birth of a child, or older adults negotiate retirement. In particular, having a confidant appears to be an especially helpful resource in coping with stress. Recent evidence highlights race, ethnicity, and culture as important factors in links between stress and social support. For example, studies suggest that social support can buffer ethnic minorities from the stress associated with discrimination.

Social support is most effective as a buffer against stress when the support matches the needs elicited by the stressor. Researchers discuss four main types of support: emotional, appraisal, informational, and instrumental. Different types of support are more appropriate as buffers against some stressful experiences than others. For example, bereavement may call for emotional support, and illness may call for instrumental support.

Researchers also distinguish between the perceived availability of support and the actual receipt of support. Available or anticipated support is the belief that social network members are willing to provide support if the need for it arises in the future. Anticipated support and received support have different effects on psychological well-being. It appears that the perception that support is available if needed has direct positive effects on well-being but that actually receiving support is not always beneficial. In particular, received support may not act as a buffer against stress because receiving support may be a signal that an individuals' own coping resources have been inadequate in resolving the stressful situation. Individuals are likely to try to deal with stress on their own before turning to others for help. Thus, receiving support may indicate that their own efforts have failed and in some sense counteract the beneficial effects of the support. In addition, receiving support may place burdens on the support network and lead to conflicts among members. Anticipated support, on the other hand, promotes individuals' own coping efforts, helps individuals maintain hope that others are available if needed, and reflects a sense of caring and commitment from the social network.

It is important to recognize that the presence of a social relationship does not necessarily indicate that the relationship is supportive. The overall quality of the relationship is more important than its mere existence. For example, Major and colleagues found that supportive relationships could help women cope with the stressful experience of having an abortion. If these supportive relationships were also perceived as being high in conflict, however, the women showed

more distress than if the relationships were low in conflict. Among women who did not perceive their relationships as supportive, there was no association between conflict in these relationships and distress. Thus, in considering whether relationships can act as buffers against stress, it is important to consider their negative as well as supportive characteristics.

To summarize, a great deal of evidence suggests that social support can help individuals cope with stress. To be most beneficial, this support should match the needs elicited by the stressor and be provided or expected in the context of a high quality relationship.

Social Support as a Source of Stress

The fact that social support can have negative consequences for the individual has only recently been addressed, and researchers are now openly calling for more attention to this issue. Social networks can become a source of conflict, stress, and sometimes health-damaging effects. One important dimension to this approach is the finding that larger social networks do not guarantee happiness, nor do they necessarily create a buffer from stress. For example, larger social networks are associated with greater stress among women, especially due to role strain. A case in point is role overload, a situation in which demands are placed on an individual that exceed his or her capacities to cope with them. For example, the responsibilities of informal caregivers to seriously impaired relatives may compete with other demands of a social network, resulting in a significant amount of stress for the individual. Having more social contacts "not only provides more potential resources but also may create additional demands on time and increase the probability of interpersonal conflicts" (Cohen and Syme, 1985: 12). Nevertheless, role theorists, such as Coontz, have noted that multiple roles are also associated with multiples sources of satisfactions, reward and accomplishments, thus suggesting the picture is not a simple one.

The most recent findings about social relations as a source of stress provide some insight into biopsychosocial links. When social relations are negative in nature, physiological reactivity becomes heightened. For instance, hostile spousal interactions increase heart rate, blood pressure, and neuroendocrine reactivity, as well as compromising immune functions. Moreover, negative social relations earlier in the life course directly affect physiological reactions later in life, contributing to allostatic load, which in turn may accelerate disease and poor mental health. This new body of research suggests that social relations influence physiological stress indicators and that, furthermore, the effects may be lifelong and cumulative.

Assistance from family is often thought to be the most reliable and consistent source of social support, yet family members who provide social support during times of crisis are themselves the sources of stress at other times. Moreover, negative interactions are more likely to involve kin than nonkin. Family also tends to make up the majority of individuals' social networks as they age. Thus, the source of stress is often found through the relationships that unfold within the confines of family support situations. For example, divorced women with social networks dominated by family may be more criticized for their failed marriage and thus experience significantly more, rather than less, stress.

Morgan conducted focus group discussions with recently widowed women and men to explore whether social networks actually ease adjustment to widowhood. The discussions revealed that a significant portion of negative outcomes derived from relationships with family members. Morgan stated that his findings may generate useful hypotheses about the role of family and friends as sources of support by suggesting that a distinction be made between positive and negative aspects of relationships. Support from family often produces stress when discrepancies exist between expectations and actual performance.

High levels of social support have also been found to cause stress. For example, research with the elderly demonstrates that such high levels of support from caregivers may cause stress and thus be harmful to the elder's well-being. Similarly, research has shown that family assistance to adolescent mothers helps them cope with many stressful aspects of their experience, but this support often is accompanied by conflict, frustration, disappointment, and stress. Too much support may foster dependency, causing the receiver to lose autonomy and hence develop low self-esteem. Another example is chronic financial strain. Stack found that survival in a poverty-ridden community depended on financial support from social network members. Support of this type, although generally positive, could under certain conditions be problematic. Having others expect support can also limit the success of the support provider. The network members they turn to for support in the face of chronic financial stress are likely to be facing financial difficulties of their own. The issue is complex. Social support can help minorities cope with the harmful mental and physical health consequences of racial discrimination. Jackson, Williams, and Torres hypothesize that, although negative health outcomes are associated with the stress of discrimination and the related experience of social inequality and racism,

these effects can be successfully buffered by social support. At the same time, it has been noted that having a strong group identity may elevate individuals' perceptions of discrimination and in some ways make them more vulnerable to its effects.

Summary and Conclusion

Many theorists from several different disciplines have developed frameworks regarding the association between social support and stress. Empirical investigations of these theories have revealed that social support can help individuals cope with stress but that social support also can be a source of stress. These theories and research suggest that social support can be a complicated matter with both positive and negative implications for individuals and their ability to cope with the stress and challenges of everyday life.

See Also the Following Articles

Caregivers, Stress and; Familial Patterns of Stress; Health and Socioeconomic Status; Sex Differences in Human Stress Response; Social Capital.

Further Reading

Antonucci, T. C. and Jackson, J. S. (1987). Social support, interpersonal efficacy, and health: a life course perspective. In: Carstensen, L. L. & Edelstein, B. A. (eds.) *Handbook of clinical gerontology*, pp. 291–311 New York: Pergamon.

Cantor, M. H. (1979). Neighbors and friends: an overlooked resource in the informal support system. *Research on Aging* **1**, 434–463.

Cohen, S. and Syme, S. L. (1985). Issues in the study and application of social support. In: Cohen, S. & Syme, S. L. (eds.) *Social support and health*, pp. 3–22 New York: Academic Press.

Coontz, S. (1992). *The way we never were: American families and the nostalgia trap*. New York: BasicBooks.

Jackson, J. S., Williams, D. R. and Torres, M. (2003). Discrimination, health and mental health: the social stress process. In: Maney, A. (ed.) *Socioeconomic conditions, stress and mental disorders: toward a new synthesis of research and public policy,* pp. 86–146. Bethesda, MD: NIH Office of Behavioral and Social Research.

Kahn, R. L. and Antonucci, T. C. (1980). Convoys over the life course: attachment, roles and social support. In: Baltes, P. B. & Brim, O. (eds.) *Life-span development and behavior* (vol. 3), pp. 254–283. Boston: Lexington.

Kiecolt-Glaser, J. K. and Newton, T. L. (2001). Marriage and health: his and hers. *Psychological Bulletin* **127**, 472–503.

Litwak, E. (1985). *Helping the elderly: the complementary roles of informal networks and formal systems.* New York: Guilford.

Major, B., Zubek, J. M., Cooper, M. L., et al. (1997). Mixed messages: implications of social conflict and social support within close relationships for adjustment to a stressful life event. *Journal of Personality and Social Psychology* **72**, 1349–1363.

McEwen, B. S. (2000). Allostasis and allostatic load: implications for neuropsychopharmacology. *Neuropsychopharmacology* **22**, 108–124.

Morgan, D. L. (1989). Adjusting to widowhood: do social networks really make it easier? *The Gerontologist* **29**, 101–107.

Pearlin, L. I. (1985). Social structure and processes of social support. In: Cohen, S. & Syme, S. L. (eds.) *Social support and health*, pp. 43–60 New York: Academic Press.

Seeman, T. E. and Crimmins, E. (2001). Social and environmental effects on health and aging: integrating epidemiologic and demographic perspectives. In: Seeman, T. E. (ed.) *Special issue on population heath and aging: strengthening the dialogue between epidemiology and demography. Annals of the New York Academy of Sciences,* 954), pp. 88–117. 88–117.

Sellers, R. M. and Shelton, J. N. (2003). The role of racial identity in perceived racial discrimination. *Journal of Personality and Social Psychology* **84**, 1079–1092.

Stack, C. B. (1974). *All our kin*. New York: Harper and Row.

Social Support in Trauma

K O'Donnell and A Steptoe
University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

Social Support and Traumatic Events
Fluctuating Social Support in Times of Crisis
Social Support and Posttraumatic Stress Disorder

Glossary

Perceived support	Perceptions of helping and support behavior by others.
Received support	The actual experience of helping and support behaviors.
Social embeddedness	The extent to which the individual has active ties with family, relatives, friends, work colleagues, and other members of the community. The size, density, and

	reciprocity of the social network are all important.
Support deterioration	The reduction in perceived support and social embeddedness that may take place following trauma and disaster.
Support mobilization	The activation of helping and support behaviors in times of trauma.

Social Support and Traumatic Events

Everyone has to deal with traumas of varying forms in his or her life. These can be personal, such as the diagnosis of a serious illness, involvement in an accident, or unexpected death of a relative; or traumas can occur on a larger scale in the community when disaster strikes. Disaster can strike unpredictably at any time and in any place. Even when disasters do not lead directly to injury, they can have devastating effects on mental and physical health, on economic well-being, and on the structure of a society. The impact of disasters depends on many factors, including the location, context, and the actual form the disaster takes. The location of a disaster affects people's access to help and the availability of resources. For example, the 2004 Indian Ocean tsunami devastated remote areas of several countries, and it took many days before assistance was received. The context of a disaster is very influential; in some cases, survivors can lose their possessions, homes, and livelihoods, with long-term consequences for health and well-being. The form of the disaster – whether it is a natural disaster, involves humanmade objects (such as trains or nuclear power plants), or is deliberate (in the case of terrorism) – also make a difference. A recent authoritative review of some 160 samples of disaster victims concluded that mental health outcomes were worse for young people, victims from developing countries (compared to those from developed countries), and victims who experienced mass violence such as terrorism or shooting sprees (compared with natural or technological disasters). At the same time, there is great variation in responses to disaster. Worse mental health outcomes and more stress symptoms are commonly observed in women (compared with men), ethnic minority groups, people exposed to more severe stress, people with other severe stresses in their lives, and people with previous psychiatric problems.

Research on how people cope with disaster and trauma and rebuild their lives has focused on the psychological and material resources that are mobilized to help them. Prominent among these is social support. It may seem obvious that social support is a good thing in situations of trauma and disaster. The news media regularly depict neighbors, friends, and even strangers rallying round when a disaster strikes. But the reality is far more complex. Often, the acute mobilization of support is not sustained in the long term, and the people who need support most are not necessarily the ones who receive it.

Social support is multidimensional. Two broad types of support can be offered in times of trauma: emotional support (comfort, sympathy, and reassurance) and practical support (food, shelter, money, advice, etc.). Social support takes place within a social and community context, and the extent to which the individual is socially embedded (has relationships within the community) is crucial. There are also important differences between received and perceived support that need to be taken into consideration. Social support is related to the individuals' socioeconomic circumstances. For example, contrary to the widely held belief that social support might be stronger in lower-socioeconomic groups due to extended ties and the values placed on social connectedness, research has shown that it is often actually lower in these groups. Mickelson and Kubzansky explain that, whereas middle-class communities can pool their resources to help those in need in times of stress, in more disadvantaged populations these resources are already stretched. Social support levels are very fluid. They depend on whether the trauma affects one individual or affects a whole social network. Studies of how social support operates during disaster and trauma has led to the recognition of the importance of concepts such as support mobilization and support deterioration.

Fluctuating Social Support in Times of Crisis

There is typically a strong mobilization of helping behavior when disaster strikes. However, the research shows that certain stressful events in people's lives can lead to a decline in social support. Divorce, the death of a loved one, the loss of a job, and illness are not only stressful in themselves but can also lead to a significant change in the availability of social contacts. There are indications that in these circumstances the decline in social contacts can itself be an additional stressor. For the individual whose partner was his or her main source of emotional support, comfort, and advice, the partner's death means that this support can no longer be called on to help in the coping process.

Similar factors may operate in disaster situations. Early research by Kaniasty and Norris was carried out with victims of floods in Kentucky in the 1980s. These disasters occurred at the same time as ongoing research that involved measures of social support, making it possible to compare support levels in the

same community both pre- and posttrauma. It was found that the victims of floods reported that less social support was available after the disaster than before. Their perceptions and expectations of support were not fulfilled. Consequently, the psychological impact of the disaster was magnified; people were not only stressed by their losses but also by the deterioration of perceived social support. This disaster took place in an impoverished rural community where there were few surplus resources.

Later research carried out in the context of two major hurricanes (Hurricane Hugo and Hurricane Andrew) broadly supports these findings. In the short-term (up to 12 months after the disasters), poor mental health is positively associated with greater direct exposure to the hurricane and is negatively related to perceived support. But perceived support was lower in people with greater direct exposure because those who were more in need were more likely to feel that support had fallen short. In the longer run, mental health was more closely associated with perceptions of support than with the extent of exposure to the disaster.

Kaniasty and Norris also examined the different ways in which social support is influenced by the context of the crisis, studying two towns in Mexico that were affected by severe floods and mudslides in 1999. The towns had very different experiences of the disaster. In one (Teziutlan), a whole mountain region was destroyed, so survivors were moved to a new community outside the original city. Flooding was worse in the second town (Villahermosa), but in this case victims were not moved. These communities were studied at several time points from 6 months to 2 years after the disaster. This allowed changes in the ratings of social support to be examined. At 6 months after the disaster, the people of Teziutlan reported less perceived support and less social embeddedness than the residents of Villahermosa. At the 2 year follow-up, the reported levels of social support in Villahermosa had improved markedly, whereas in Teziutlan they remained below normal. Social support was particularly low among women and people of lower social status.

Social Support and Posttraumatic Stress Disorder

The consensus in the literature is that social embeddedness and received support are beneficial for the mental health of disaster victims. Lack of social support is also a strong risk factor for posttraumatic stress disorder (PTSD), an effect that has been observed across a range of study populations. Interestingly, it appears that the level of social support available at the time of the trauma and after the trauma is crucial to the development of PTSD symptoms. The levels of social support measured before the trauma are less predictive, probably because perceptions of support may not match reality. But relationships can be complicated. One study of police officers in New Zealand examined the effect of support from peers, from supervisors and from people outside work on the development of PTSD symptoms following exposure to severe trauma. Although levels of support did help to explain some of the variation in PTSD, it was found that being able to talk to others about the negative aspects of the work was actually associated with more, rather than fewer, PTSD symptoms. Fostering the right atmosphere for social support in organizations in which staff may have to deal with high levels of trauma can be of benefit.

The extent to which the observational research on social support and trauma can be translated into active support services by health professionals is controversial. The fact that the people who report low support from their friends and families experience greater psychological distress than others does not mean that remedial support by health and emergency personnel will necessarily be beneficial. There is growing evidence, for example, that some early psychological interventions such as debriefing can do more harm than good. Similarly, research on patient support groups indicates that these can have deleterious as well as beneficial effects. Victims of disasters can be reluctant to seek or accept psychological help. It may be that social support interventions are best delayed for a number of weeks after a disaster, prioritizing more acute needs. It is only after the disaster situation has settled down to some extent that victims are prepared to focus on emotional needs. Further research is needed to extend knowledge of how communities affected by disaster could be supported to aid the sustained mobilization of social support, whether support interventions are possible, and which approaches are most beneficial for those at risk.

See Also the Following Articles

Combat, Acute Reactions to; Disaster Syndrome; Disasters and Mass Violence, Public, Effects of; Earthquakes, Stress Effects of; Floods, Stress Effects of; Social Support; 9/11, Religion and Stress; Crisis Intervention; Loss Trauma; Posttraumatic Stress Disorder – Clinical.

Further Reading

Brewin, C. R., Andrews, B. and Valentine, J. D. (2000). Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. *Journal of Consulting and Clinical Psychology* **68**, 748–766.

Galea, S., Ahern, J., Resnick, H., et al. (2002). Psychological sequelae of the September 11 terrorist attacks in New York City. *New England Journal of Medicine* **346**, 982–987.

Helgeson, V. S. and Cohen, S. (1996). Social support and adjustment to cancer: reconciling descriptive, correlational and interventional research. *Health Psychology* **15**, 135–148.

Kaniasty, K. and Norris, F. H. (1993). A test of the social support deterioration model in the context of natural disaster. *Journal of Personality and Social Psychology* **64**, 395–408.

Mickelson, K. D. and Kubzansky, L. D. (2003). Social distribution of social support: the mediating role of life events. *American Journal of Community Psychology* **32**, 265–281.

Norris, F. H., Friedman, M. J., Watson, P. J., et al. (2002). 60,000 disaster victims speak. Part I: An empirical review of the empirical literature, 1981–2001. *Psychiatry* **65**, 207–239.

Norris, F. H., Baker, C. K., Murphy, A. D., et al. (2005). Social support mobilization and deterioration after Mexico's 1999 flood: effects of context, gender and time. *American Journal of Community Psychology* **36**, 15–28.

Stephens, C., Long, N. and Miller, I. (1997). The impact of trauma and social support on posttraumatic stress disorder. A study of New Zealand police officers. *Journal of Criminal Justice* **25**, 303–314.

Somatic Disorders

F Creed
University of Manchester, Manchester, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F Creed, volume 3, pp 483–485, © 2000, Elsevier Inc.

Psychosomatic Disorders
Stress and the Onset of Organic Disorders
Medically Unexplained Symptoms
Anxiety and Depressive Disorders

Glossary

Anxiety disorder	A disorder of mood characterized by tension and worry, often accompanied by physical symptoms of tension, including palpitations, butterflies in the stomach, and breathing difficulties.
Chronic social difficulties	Ongoing environmental stressors such as unemployment, serious financial or health difficulties, and ongoing serious difficulties in close relationship.
Depressive disorder	A disorder of mood accompanied by sleep, appetite disturbance, lack of energy and concentration, and possibly ideas of suicide.
Life event	A discrete change in the person's immediate environment, such as a bereavement or a marital separation.
Medically unexplained symptoms	Symptoms of pain or function (e.g., bowel or breathing disturbances) that cannot be explained on the basis of underlying organic disease.
Organic disease	A physical disease understood on the basis of recognized tissue damage.
Psychosomatic medicine	The study of biological, psychological, and social aspects of all diseases.

Psychosomatic Disorders

Psychosomatic disorders were previously thought to be a separate group of disorders in which stress and psychological distress led to disease. Examples were peptic ulcer, arthritis, and dermatitis. It is now recognized that all diseases may have social, psychological, and behavioral aspects as well as physical organic features. The World Health Organization defines psychosomatic medicine as "the study of biological, psychological and social variables in health and disease."

This is best illustrated by the example of heart disease. Heart disease is caused by many factors. These include inherent biological factors, including age, sex, and genetic enhancement, that increase the chances of a heart attack; physiological factors, such as raised blood pressure, smoking, alcohol consumption, and increased blood cholesterol level; and psychological and social factors, which include depression, anxiety, and certain forms of stress, which are associated with an increased chance of developing a heart attack or increased chance of dying thereafter. Stress may lead to anxiety or depressive disorders, which can lead to a fast heart rate and increased chance of irregular heart beats, which may be fatal soon after a heart attack. Lack of social support – in the form a close person with whom all difficulties or problems can be shared – also contributes to increased chance of further heart attacks.

Thus, the modern view of psychosomatic mechanisms includes a role for stress, depression, and lack of social support alongside biological factors in the causation or outcome of disease. The relative importance of environmental stress varies in different conditions.

Stress and the Onset of Organic Disorders

Stress is much less important than the other biological or physical causative factors in the development of a heart attack. There is no clear link between stressful life events (e.g., marital separation and loss of job) and having a heart attack. There is a link with more chronic stress – 40% of heart attack patients have experienced a chronic social difficulty concerning work, money, or housing compared to 17% of healthy controls. The greatest difference seems to be in the excessive hours worked and lack of proper holidays. This seems to be something that individuals bring on themselves by working excessively hard (i.e., the type A personality).

In other diseases, different patterns have been reported. In peptic ulcer, for example, life events are important, especially those that engender goal frustration. These events involve working hard toward a particular goal for many years and then being thwarted by a change of circumstances at the last moment. Such an event occurred in over half (54%) of people with peptic ulcer, compared to only 9% of healthy comparison groups.

Another stress is that of a major earthquake. There was an increase in the number of people with stomach ulcer after the Hanshin-Awaji earthquake; for most people, the ulcer developed in conjunction with an infectious agent in the stomach, *Helicobacter pylori*, but among those physically injured following the earthquake stomach ulcers developed independently of this infectious agent. Thus, there is a clear relationship with stress, which must be considered alongside other risk factors for peptic ulcer – *H. pylori* infection, drugs taken for arthritis smoking, and an inherited predisposition.

There is a discernible increase in environmental stress prior to the onset of stroke, but the effect is weak – the biological factors listed as risk factors for heart disease are much more important. In one study, 24% of stroke patients and 12% of controls had experienced a severe life event in the year prior to the stroke. This is a much smaller difference than that found in peptic ulcer.

It is easy to assess life events prior to the onset prior of a heart attack or a stroke because these conditions develop at a clear point in time. In multiple sclerosis, however, the onset is much less clearly defined. It does appear that the experience of a severe life event during the 6 months prior to the onset of multiple sclerosis is very considerable – 62% of multiple sclerosis patients had experienced such an event, compared to only 15% of a control group. If this result is confirmed, it suggests that the experience of stress probably interacts with the immune system, which may have been sensitized by exposure to an infection during childhood or early adulthood.

Stress and the Immune System

One of the most likely ways that stress leads to the onset of organic disease is by the modification of the immune system. This has been shown in very well-controlled studies of susceptibility to infection with the cold virus. An impressive feature is the dose–response relationship between an index of psychological stress and chance of infection; that is, the more stress experienced, the greater likelihood that a cold will develop after experimental exposure to the virus. The psychological stress in these studies comprised a combination of stressful life events (environmental stress) and the degree of distress (symptoms of anxiety and depression).

In order to exclude the interference with biological variables, these studies controlled for a number of variables that might explain the stress–infection association: smoking, alcohol, exercise, diet, quality of sleep, number of white cells in the blood, and total immunoglobular levels. The researchers also excluded the effect of personality variables of low self-esteem, personal control, and degree of extroversion and introversion. When this is done, there is a clear link between the stress levels of the caregivers of people with Alzheimer's disease and immune response, measured by the antibody response to influenza vaccination.

Medically Unexplained Symptoms

By contrast to the pattern observed in organic disease, there seems to be a very clear increase in stress just before the onset of certain disorders in which no organic disease can be found to explain the symptom. Typical disorders are irritable bowel syndrome, chronic fatigue syndrome, and fibroymalgia. The symptoms may be severe (e.g., pain) and may lead to an operation, such as an appendectomy although the appendix appears to be quite normal The pain may also lead to many investigations that fail to reveal any recognized organic disease to explain it. Approximately two-thirds of such patients have experienced recent environmental stress, compared to less than one-quarter of people with organic disease

or healthy control subjects. This pattern of recent stress is similar to the stress that occurs prior to depressive disorders.

Anxiety and Depressive Disorders

The most common reaction to environmental stress is depression or anxiety, which, if severe and/or prolonged, amounts to anxiety or depressive disorder. Bereavement is a typical environmental stress that may be followed by the onset of depression; it may also be associated with the increased onset of serious physical illness. The mortality of widowers has been shown to increase 40% during the 6 months following the loss of their wives.

This means that the combination of a physical (organic) disorder and concurrent anxiety or depressive disorder is common; approximately one in five people with a serious and/or disabling physical illness have concurrent anxiety or depression. Depressed people with physical illness experience more severe pain, view their illness more seriously, and feel that the chance of recovery is lower. This view of the illness can lead to a delay in the return to normal activities, such as return to work or resuming a normal sexual relationship after a heart attack.

A woman with rheumatoid arthritis, for example, whose husband also has a disabling illness can be described as being under stress, partly because of her own illness and partly because of her husband's illness. These two stresses may lead to depression, in which case the pain worsens and the illness appears to be worse. Her going to a rheumatology clinic may lead to a change in the treatment for her rheumatoid arthritis, but only if the depression is recognized and treated will the patient's health begin to return. For this reason, modern medicine is increasingly keen to identify anxiety and depressive disorders in patients with chronic diseases, as well as the accompanying environmental stress.

Further Reading

Creed, F. H. (1993). Stress and psychosomatic disorders. In: Goldberger, L. & Breznitz, S. (eds.) *Handbook of stress: theoretical and clinical aspects*, pp. 496–510. New York: Macmillan.

Cohen, S., Tyrrel, D. A. J. and Smith, S. P. (1991). Psychological stress and susceptibility to the common cold. *New England Journal of Medicine* **325**, 606–612.

Dickens, C. M., McGowan, L., Percival, C., et al. (2004). Lack of a close confidant, but not depression, predicts further cardiac events after myocardial infarction. *Heart* **90**(5), 518–522.

Frasure-Smith, N., Lesperance, F. and Talajic, M. (1993). Depression following myocardial infarction: impact on 6-month survival. *Journal of the American Medical Association* **270**, 1819–1825.

Hemingway, H. and Marmot, M. (1993). Evidence based cardiology: psychosocial factors in the aetiology and prognosis of coronary heart disease. Systematic review of prospective cohort studies. *British Medical Journal* **318**, 1460–1467.

Matsushima, Y., Aoyama, N., Fukuda, H., et al. (1999). Gastric ulcer formation after the Hanshin-Awaji earthquake: a case study of Helicobacter pylori infection and stress-induced gastric ulcers. *Helicobacter* **4**, 94–99.

Vedhara, K., Cox, N. K., Wilcock, G. K., et al. (1999). Chronic stress in elderly carers of dementia patients and antibody response to influenza vaccination. *Lancet* **353**, 627–631.

World Health Organisation (1964) *Psychosomatic Disorders. Thirteenth Report of the WHO Expert Committee on Mental Health*. Geneva: World Health Organisation.

Space Travel *See:* Pressures, Effects of Extreme High and Low; Space, Health Risks of.

Space, Health Risks of

V A Convertino
U.S. Army Institute of Surgical Research, Fort Sam Houston, TX, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction
Stress Environment of Space
Impact on Body Fluid and Electrolyte Balance
Impact on Nutrition
Impact on the Cardiovascular System
Impact on Thermoregulation
Impact on the Skeletal Muscle System
Impact on the Skeletal System
Impact on Neurovestibular Function
Impact on the Immune System
Summary

Glossary

Atrophy	A diminution or degeneration of bodily tissues or organs.
Baroreceptors	Neural structures located throughout the cardiovascular system that transmit nerve signals to the central nervous system in response to changes in blood pressure and initiate an involuntary nerve response (baroreflex) that brings arterial blood pressure back toward normal.
Cardiac output	The quantity of blood pumped by the heart each minute; the product of the amount of blood pumped each time the heart contracts (stroke volume) and the number of times the heart beats (heart rate).
Hydrostatic pressure	The pressure created from the weight of a column of fluid due to the presence of gravity. On Earth, hydrostatic pressure increases 1 mmHg for each 13.6 mm distance below the surface of the fluid column.
Ionizing radiation	Energy imparted by radiation to an atom that results in the ejection of an electron and the creation of an ion.
Maximal O_2 uptake	The maximal capacity for consuming and using oxygen (aerobic capacity) by the body tissues.
Muscle fiber	An individual muscle cell.
Orthostatic intolerance	The inability to tolerate the upright posture due to low blood pressure and reduced blood flow to the brain, leading to symptoms such as lightheadedness, dizziness, blurred vision, and possible fainting.
Stress	A challenge to homeostasis that is accompanied by a change in physiological function.

Introduction

Human physiology and health have evolved and adapted in the presence of Earth's gravity. The physical stress imposed by movement in the presence of gravity manifests itself in a dramatic elevation in energy exchange, reflected by such physiological responses as increased oxygen uptake, respiration, heart rate, and sweating. The magnitude of physical stress associated with the development of the force required by muscles to perform various physical activities such as moving one's body or lifting or throwing objects is dependent on the level of gravity opposing these actions. Thus, gravity has traditionally been considered the primary constant stress factor for influencing adaptation and maintaining the health of the human body on Earth. However, the evolution of travel and habitation of humans in space, and the associated absence of gravity, has imposed new physiological adaptations via a reduction in hydrostatic pressure gradients within the cardiovascular system, a decreased weight load on the muscles and bones, and a lower energy requirement and/or output. The space environment is unique in that the physiological stress normally associated with the maintenance of health on Earth is altered as a result of reduced gravitational accelerations and factors related to confinement in a closed ecological environment. This article describes numerous physiological adaptations and responses associated with travel to space and planets with lower gravitational forces than Earth that can compromise normal function and health, particularly on return to Earth.

Stress Environment of Space

Space exposes astronauts to extreme environmental and psychological conditions imposed by a variety of stressors that result in a pronounced compromise to their safety and health, both in the space environment and on return to Earth. Immediately on entry into orbit, the relative absence of gravity results in unloading of the body tissues; external light–dark cycles occur every 90 min, disrupting normal circadian rhythmicity and ultraviolet light is nearly abolished;

microflora proliferate in the air and water and carbon dioxide levels within the enclosed space craft are chronically elevated. The crew members are required to exercise vigorously during long-duration space missions, often more intensely than they did before flight in an effort to maintain their baseline fitness and reduce the physiological decompensation associated with space flight. Psychological stress levels that can lead to significant neuroendocrine effects on physiological functions may become elevated when diverse crew members are required to exist in relatively confined living quarters, with high noise and vibration levels and with altered sleep and eating patterns. Indeed, there are reports suggesting that interpersonal conflict and depression may have contributed to early mission termination. Perhaps the most threatening factor to the health of individuals during extended travel in space is the significant elevation in ionizing radiation. For instance, the dose of ionizing radiation received during a 6-month stay on the International Space Station can be 15-fold greater than living 1 year in Houston, Texas. Depending on the type and dose of exposure, ionizing radiation can lead to damaged nerve cells in the central nervous system, cataracts, reduced fertility, and potential carcinogenic hazards. Hence, it is clearly a misconception to view space as a condition of reduced stress. The physiological responses to any one or a combination of the stressors associated with space missions can result in a loss in function and a subsequent compromise of health, particularly on return to a higher-gravity environment such as Earth.

Impact on Body Fluid and Electrolyte Balance

In the absence of gravity, altered hydrostatic pressure gradients within the cardiovascular system represent a potent stressor experienced immediately on exposure to space, resulting in a redistribution of body fluids that increases perfusion in the upper body regions while decreasing perfusion in the lower body regions. Increased filling of the heart and central circulation activate volume receptors that initiate an inhibition of vasopressin (a water-retaining hormone) and the renin-angiotensin-aldostersone system (a salt-retaining hormone response) that regulate body fluid and electrolyte balance. These neuroendocrine responses result in the increased excretion of water (diuresis) and sodium (natriuresis) from the kidneys. Although there may not be clinically significant alterations in electrolyte concentrations in the body fluids, plasma volume is reduced within 24–48 h of exposure to space. The resulting elevation in hemotocrit (ratio of red blood cells to total blood volume) inhibits the secretion of a hormone responsible for stimulating red blood cell production (erythropoietin) and subsequently induces a more gradual decrease in red blood cell mass. The result of kidney excretion of sodium chloride and reductions in plasma and red blood cell volumes is a resetting of the total circulating blood volume by 10–15% below the preflight level with little change in the primary electrolyte and protein concentrations in the body fluids.

Impact on Nutrition

In addition to the involuntary loss of body fluids and electrolytes, space alters dietary intake through changes in eating behaviors. The most prominent evidence that minimal nutritional requirements are compromised in space is a significant average reduction of approximately 7% ($\sim$4 kg) in body weight. However, the average nutritional composition of the typical diet during space missions is similar to that on Earth with $\sim$16–18% protein, $\sim$17–25% fat, $\sim$60% carbohydrate, $\sim$1% fiber, and $\sim$3% minerals. Average daily fluid ($>$2000 ml) and sodium ($>$3500 mg) intakes are at or above recommended minimal requirements, whereas the average dietary calcium intake of approximately 900 mg day^{-1} is below the recommended dietary allowance (RDA) of 1000–1200 mg day^{-1}. The average daily baseline energy expenditure in space has been estimated at 2500 kcal, although the actual daily energy expenditure is much often higher than this because of the performance of physical exercise and infrequent extravehicular (outside the spacecraft) work activities (EVA). When the data are standardized for body size (weight), it becomes clear that space inhibits hunger because there is a significant and consistent reduction in daily energy intake and a negative nitrogen balance. Reduced caloric intake (i.e., failure to consume available food) in space has been associated with crewmembers' decreased appetite, feeling of fullness in the abdomen, nausea, and preoccupations with critical mission tasks.

Impact on the Cardiovascular System

The loss of circulating blood volume contributes directly to a reduction in cardiac filling, as evidenced by the reduction in echocardiographic measures of end-diastolic (input) and stroke volumes (output) in astronauts following space flight. Despite compensatory elevations in heart rate, a subsequent reduction in cardiac output can result in low blood pressure (hypotension) if compensatory peripheral vasoconstriction is not adequate. In addition to reduced circulating vascular volume, there is evidence of

alterations in cardiovascular structures that occur during adaptation to space. Magnetic resonance imaging of cardiac muscle mass before and after space missions has revealed evidence of a smaller myocardial mass (cardiac atrophy), which could contribute to the reduced myocardial contractility and stroke volume that is observed when astronauts return to Earth from space. In the relative absence of gravity and posture-related hydrostatic pressure gradients, the space environment elicits much smaller fluctuations in pressures within the cardiovascular system than those that naturally occur during change in posture in the terrestrial gravity environment. Without the dynamic stimulation of the baroreceptors (pressure receptors) from such pressure fluctuations, space is associated with a compromise to reflexes that regulate blood pressure and flow (baroreflexes) only on return to Earth. There is also evidence of elevated sympathetic nerve activity and circulating catecholamines during exposure to the space environment. Higher sympathetic nerve activity in the presence of lower blood volume is associated with the elevated constriction of the arteries in the peripheral circulation, a condition that represents a lower capacity to maintain blood pressure during standing. The condition of less circulating blood volume can be exacerbated by an increased venous compliance of the lower extremities that can lead to a greater accumulation of blood away from the brain and heart and has also been reported following exposure to space.

There are at least two documented functional compromises to the health and safety of astronauts that are associated with these cardiovascular alterations during space. On return to Earth, some degree of cardiovascular instability due to the development of low arterial blood pressure during standing (orthostatic hypotension) has been reported in many individuals who have traveled in space. Significant orthostatic hypotension has been described in 40–50% of crew members in the U.S. space program and is associated with symptoms such as dizziness, lightheadedness, and blurred vision. In some astronauts, this condition has led to fainting (syncope) and is defined as orthostatic intolerance. Orthostatic hypotension and intolerance after return from space are associated with reduced cardiac output, which is associated with low circulating blood volume, which is further complicated by the increased pooling of blood in the lower extremities due to increased venous compliance (extension of the veins) and attenuated elevation in heart rate due to cardiac baroreflex impairment. Compromised orthostatic performance in astronauts is of particular operational interest because of its potential to increase the risk of serious injury from falling, particularly during critical times such as an emergency egress from the spacecraft. This potential risk has raised great concern and, subsequently, led to vigorous research to counteract the problem with various countermeasures.

On return from space, astronauts experience more than a 20% reduction in their maximal oxygen uptake (VO_2max; aerobic capacity). The measurement of VO_2max provides an integrated index of the functional integrity of the cardiorespiratory system and its mechanisms that regulate the delivery to and use of oxygen by the working muscles. The reduction in VO_2max on return to Earth can be explained entirely by a decrease in cardiac filling and stroke volume associated with lower circulating blood volume. The cardiovascular effect of the reduction in VO_2max on return to Earth is highlighted if physical work is performed in the upright (i.e., orthostatic) posture, as evidenced by higher heart rates and lower stroke volumes measured in astronauts who exercise in the upright compared to the supine posture after their return from space. The operational consequence of reduced aerobic reserve is an increased stress on the cardiovascular system with any physical effort during or after return from a space mission. Although there is no documented case of an adverse cardiac event on return to Earth after a space mission, a potential health consequence of increased stress on the heart, particularly during intense exercise or physical work, could manifest itself in myocardial ischemia or abnormal cardiac rhythms.

Impact on Thermoregulation

On Earth, the body core temperature fluctuates within $\pm 0.5^{\circ}$C as a consequence of circadian rhythm, with higher temperatures during the day than the night. Body temperature circadian rhythm is altered in space such that the minimum night temperature is elevated by 0.2°C and the maximal daily temperature is reduced by 0.2°C. Despite a change in the natural light–dark cycles to 90 minutes, the change in body temperature is in part a direct effect of low gravity because body temperature circadian rhythm can be altered by changes in gravity in the presence of normal light–dark cycles. While in space, the average sweat rate is 10% less and insensible water exchange is lower during the performance of physical work because convective heat flow and sweat drippage are reduced in the low-gravity environment by the formation of an observed sweat film on the skin surface. This results in high levels of sustained skin wetness that acts to suppress sweating. Consequently, the average sweat rate during exercise is 10% less in space and crew members demonstrate a faster rise in body core temperature for a given absolute level of

physical work after return from space. After return to Earth, skin blood flow is reduced at rest (probably associated with elevated sympathetic vasoconstriction) and the sensitivity of active vasodilation and sweating are impaired during exercise. Although there are few measurements of body temperatures recorded on astronauts during their daily activities in space, recorded alterations in the response to body temperature during daily exercise suggest a change in thermoregulation that could compromise the capacity for heat exchange and lead to higher thermal loads during EVA. The functional consequence of alterations in thermoregulation following space flight is an impaired ability to dissipate body heat during physical work, with a higher risk of heat injury.

Impact on the Skeletal Muscle System

A significant stress associated with the relative absence of gravity in space is the unloading and subsequent disuse atrophy of skeletal muscles, especially those of the lower extremities. The magnitude of the resulting decrease in limb circumference has been used as a gross index of muscle atrophy. In addition, muscle biopsies taken from astronauts before and after space missions have revealed significant muscle atrophy as evidenced by reduction in cross-sectional areas (CSAs) of slow-twitch and fast-twitch muscle fibers of the vastus lateralis (a quadriceps muscle). The relative reduction in muscle fiber CSAs is approximately 50–60% greater in larger fast-twitch fibers, which are associated with force development (i.e., strength), than in smaller slow-twitch fibers, which are associated with posture and endurance. Muscle atrophy in space is also accompanied by reduced capillary-to-fiber ratio in both slow- and fast-twitch fibers after flight, a condition that can contribute to lower oxygen delivery and use and the cardiovascular effects of limiting VO_2max.

In addition to changes in muscle morphology, spaceflight causes histochemical alterations. Although there are no significant changes in anaerobic (glycolytic) enzyme activities in muscles following space flight, the activities of enzymes associated with aerobic metabolic pathways (succinate dehydrogenase, citrate synthase, β-hydroxyacyl-CoA dehydrogenase) and the use of oxygen are reduced in both slow- and fast-twitch muscle fibers. Thus, prolonged muscle unloading and reduced energy requirements associated with space causes muscle atrophy with subsequent compromises in the capacity to generate force and in the delivery and utilization of oxygen at the level of muscle fibers.

The alterations in cellular morphology and histochemistry of skeletal muscle resulting from space travel are associated with detriments in function. Muscle strength can be decreased by 20% for knee extensors and 10% for knee flexors. In addition to the loss in muscle strength, an increase in muscle fatigue is evidenced by a 50% increase in motor nerve activity (i.e., amplitude of electromyographic recordings) required to develop an equal force in calf skeletal muscles following a space mission compared with preflight measures.

The reductions in the ability to develop and maintain dynamic muscle forces that are required for normal daily functions on Earth represents a clinical consequence of muscle atrophy associated with space. For instance, a 10% decrease in force generation in the muscle groups of the lower extremities associated with adaptation to space is associated with a 10–15% reduction in distances of long and high jumps. The atrophy and fatigue of postural muscles can limit the ability to function during simple standing. The significant loss of muscle structure and function could predispose individuals to serious injury during physical activities or render them incapable of performing routine tasks required for normal living such as standing, stair climbing, and getting out of bed.

Impact on the Skeletal System

The result of the space environment on the integrity of the skeletal structure is clear; bone mass and bone mineral density (BMD) decrease in space. The most profound cause of bone loss in space is the dramatic unloading of bone with the relative absence of gravitational forces and dynamic mechanical forces exerted by skeletal muscles during movement on Earth. Bone loss begins immediately on unloading and is associated with a rapid and sustained loss of calcium from the body. The state of the skeletal system is dictated by bone mass turnover, which is determined by the net balance between resorption and formation. Bone resorption increases within the first few days of space, with an overall increase in bone resorption of 50% during long-duration flights, as indicated by elevated urinary calcium excretion. Although not all data support the notion of changes in bone formation in space, there exists some biochemical evidence that suggests a possible reduction in bone formation such as slight reductions in osteocalcin and bone-specific alkaline phosphatase. Thus, it appears that bone loss in space is due to increased bone resorption compounded by a slowly developing decrease in bone formation. Bone calcium can be lost at rates of up to 140 mg day^{-1} during flight. During 4–14 months in space, whole body BMD decreases at a rate of approximately 0.6% per

month despite the implementation of physical exercise countermeasures of 1–2 h day^{-1}.

In addition to the unloading of bone with the absence of gravitational forces and dynamic mechanical forces exerted by skeletal muscles, other characteristics of the space environment are not conducive to the maintenance of a healthy skeletal structure. The near abolishment of ultraviolet light within the spacecraft environment can contribute to the reported reduction in vitamin D required for normal bone formation and breakdown (resorption), leading to the inhibition of calcium absorption in the intestine, parathyroid hormone release, and mineralization of bone. Ambient CO_2 levels, which are normally 0.03% on Earth, are chronically elevated to approximately 0.4% inside the spacecraft during nominal crew activity, with peaks of 1–3% during physical exercise workouts or visits from additional crew members. These high ambient CO_2 levels in the spacecraft can result in a more acid state within the body tissues, requiring bone resorption (bone breakdown at the expense of maintaining bone density) to provide bases such as carbonates and phosphates to buffer the respiratory acidosis. Although not consistently reported in all astronauts, the psychological stress of space travel may also contribute to bone loss by the direct effect that increased cortisol (stress hormone) has on stimulating bone resorption. In addition, abnormal nutritional habits reported in some crew members, such as reduced energy and calcium intake and a high sodium diet, may also contribute to the increased bone loss in space.

Health risks associated with bone degradation in space have been proposed by some flight surgeons. The increased blood and urinary calcium concentrations can increase the risk of kidney stones. There are medical reports of the formation of kidney stones in space that have resulted in infection and discomfort that could threaten the success of a space mission. Because the qualitative changes in bone in space are similar to those reported in patients with disuse osteoporosis, an increased risk of bone fracture on return to Earth represents one of the greatest health concerns of long exposures to space.

Impact on Neurovestibular Function

Neurovestibular function is important for maintaining normal posture and balance on Earth. There are several anecdotal reports by astronauts that indicate significant impairment to neurovestibular function after returning from space. While driving a car on a curved ramp, one astronaut voluntarily pulled off the road because he was experiencing a strong feeling of tilt. In another instance, an astronaut fell to the ground while attempting to walk across a room to his bed in the dark. The gravitational unloading of the vestibular system in space causes numerous adaptations that can compromise the postural stability of an individual on return to Earth. Astronauts with their eyes closed significantly overestimate the degree of tilt when placed in a chair that is rotated sideways or tilted in a roll direction after returning from space. Experiments following space flight clearly demonstrate that astronauts show the greatest postural instability when they must rely on vestibular rather than visual information to maintain posture. Because of its importance in normal daily functions such as walking, running, navigation, and driving a car, impaired neurovestibular function associated with adaptation to space may increase the risk of injury from accidents and falls following return to Earth. Fortunately, most data indicate rapid improvement in balance and posture within hours of returning to Earth.

Impact on the Immune System

There is strong evidence that humans who fly in space have compromised immune function. During the Apollo missions, 15 of 29 astronauts reported bacterial or viral infections either during flight or within the first week after returning to Earth. During the infamous Apollo 13 mission, Fred Haze suffered from a severe urinary tract infection. The examination of cell-mediated immune responses measured from blood samples drawn from various astronauts and cosmonauts after their return from space revealed a lower immune response as indicated by an inhibited mitogen-induced formation of leukocytes, a severe decrease in the ability of leukocytes to produce interferon-α/β, an altered leukocyte subset distribution and interferon production, and reduced natural killer cell activity. Although the cause(s) of compromised immune function in space is unclear, the stress associated with confined environments, intense work schedules, exposure to radiation, poor nutrition, sleep deprivation, and/or the absence of gravity have all been associated with immune suppression. In space, more frequent illnesses associated with bacterial or viral infections can compromise the successful completion of a space mission by reducing the capacity to perform physical work. When back on Earth, a significant risk to the health of a space traveler is that immune dysfunction might make the crew more susceptible to more chronic tissue radiation damage and to the future development of tumors and cancers.

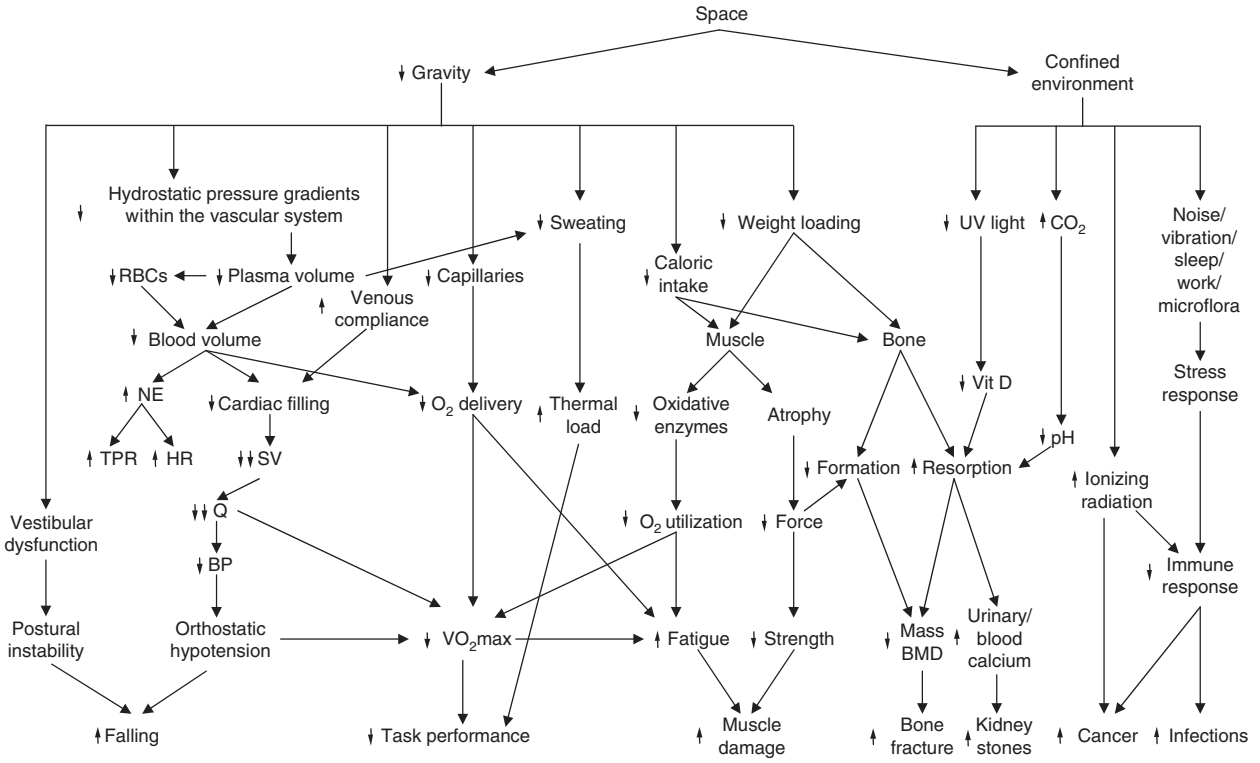


Figure 1 Cascade diagram of the relationships between physical and physiological stresses associated with the space environment and alterations in physiological functions associated with compromised health status. BMD, bone mineral density; BP, blood pressure; HR, heart rate; NE, norepinephrine; Q, cardiac output; RBCs, red blood cells; SV, stroke volume; TPR, total peripheral resistance; UV, ultraviolet; Vit D, vitamin D; VO_2max, maximal oxygen uptake.

Summary

The reduction of gravity in space presents an environmental stress that has profound effects on the structure and function of virtually all human tissues, resulting in physiological responses often seen in conditions of compromised health associated with aging, deconditioning, or inactivity-related diseases. The cascade of physical, physiological, and psychological responses associated with exposure to space is complex and interactive (Figure 1).

Although the absence of gravity is probably the most significant cause of physiological adaptations in space, the confined environment of space travel with changes in light, ambient gases, ionizing radiation levels, bacterial growth, and diverse personalities living together in confined quarters for long periods of time in the presence of high noise and vibration levels, intense work schedules, and altered sleep and eating patterns can contribute to organ dysfunctions. The space environment presents numerous physical and chemical conditions that can have profound effects on physiological health, the most significant effects of which occur on return to Earth's gravity.

Further Reading

Buckey, J. C. (2006). *Space physiology.* New York: Oxford University Press.

Buckey, J. C., Lane, L. D., Levine, B. D., et al. (1996). Orthostatic intolerance after spaceflight. *Journal of Applied Physiology* **81**, 7–18.

Convertino, V. A. (1990). Physiological adaptations to weightlessness: effects on exercise and work performance. *Exercise and Sports Science Reviews* **18**, 119–165.

Convertino, V. A. (1991). Neuromuscular aspects in development of exercise countermeasures. *Physiologist* **34** (supplement), S125–S128.

Convertino, V. A. (1996). Clinical aspects of the control of plasma volume at microgravity and during return to one gravity. *Medicine and Science in Sports and Exercise* **28**, S45–S52.

Convertino, V. A. (1996). Exercise and adaptation to microgravity environments. In: Fregly, M. J. & Blatteis, C. M. (eds.) *Handbook of physiology. Sec. 4: Environmental physiology. III: The gravitational environment*, pp. 815–843. New York: Oxford University Press.

Convertino, V. A. (1997). Conditions of reduced gravity. In: Low, P. A. (ed.) *Clinical autonomic disorders* (2nd edn., pp. 429–440). Philadelphia: Lippincott-Raven.

Convertino, V. A. (2000). Effects of microgravity on exercise performance. In: Garrett, W. E. Jr. & Kirkendall, D. (eds.) *Exercise and sport science*, pp. 459–469. Philadelphia: Lippincott Williams & Wilkins.

Convertino, V. A. (2002). Mechanisms of microgravity-induced orthostatic intolerance and implications of effective countermeasures: overview and future directions. *Journal of Gravitational Physiology* **9**, 1–12.

Convertino, V. A. (2002). Planning strategies for exercise and nutrition countermeasures for long duration space flight. *Nutrition* **18**, 880–888.

Edgerton, V. R. and Roy, R. R. (2000). Gravitational biology of the neuromotor systems: a perspective to the next era (invited rev.). *Journal of Applied Physiology* **89**, 1224–1231.

Edgerton, V. R., Zhou, M-Y., Ohira, Y., et al. (1995). Human fiber size and enzymatic properties after 5 and 11 days of spaceflight. *Journal of Applied Physiology* **78**, 1733–1739.

Fitts, R. H., Riley, D. R. and Widrick, J. J. (2000). Microgravity and skeletal muscle (invited rev.). *Journal of Applied Physiology* **89**, 823–839.

Lesnyak, A., Sonnenfeld, G., Avery, L., et al. (1996). Effect of SLS-2 spaceflight on immunologic parameters of rats. *Journal of Applied Physiology* **81**, 178–182.

Levine, B. D., Lane, L. D., Watenpaugh, D. E., et al. (1996). Maximal exercise performance after adaptation to microgravity. *Journal of Applied Physiology* **81**, 686–694.

Oman, C. M., Pouliot, C. F. and Natapoff, A. (1996). Horizontal angular VOR changes in orbital and parabolic flight: human neurovestibular studies on SLS-2. *Journal of Applied Physiology* **81**, 69–81.

Prisk, G. K. (2000). Microgravity and the lung (invited rev.). *Journal of Applied Physiology* **89**, 385–396.

Schneider, S. M. and Convertino, V. A. (2005). Physiological systems and their responses to conditions of microgravity and bed rest. In: Tipton, C. M. (ed.) *ACSM's advanced exercise physiology*, pp. 595–619. Baltimore: Lippincott Williams & Wilkins.

Stein, T. P., Leskiw, M. J. and Schluter, M. D. (1996). Diet and nitrogen metabolism during spaceflight on the shuttle. *Journal of Applied Physiology* **81**, 82–97.

Turner, R. T. (2000). What do we know about the effects of spaceflight on bone? (invited rev.). *Journal of Applied Physiology* **89**, 840–847.

West, J. B. (2000). Historical perspectives: physiology in microgravity. *Journal of Applied Physiology* **89**, 379–384.

Spinal Cord Injury, Physical Stress of

S Bhatia and M Quigley
Alllegheny General Hospital, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Cardiovascular Complications
Respiratory Complications
Gastrointestinal Complications
Urologic Complications
Skin Complications
Bone and Joint Complications

Glossary

Autonomic nervous system	Part of the nervous system that regulates several vital functions through two mutually opposing systems, the sympathetic system and the parasympathetic system.
Bradycardia	A slow heart rate, usually less than 60 beats per minute in adults.
Cauda equina	A bundle of spinal nerve roots below the termination of the spinal cord at the first lumbar vertebra. Typically it is more resistant to injury than the cord itself.
Conus medullaris	The tapering terminal part of the spinal cord at the level of the first lumbar vertebra that contains the nerves that serve distal lower extremity function and sphincter control.
Detrusor	The muscle present in the bladder wall that contracts and pushes urine out during micturition.
Ileus	Loss of muscular contractions in the bowel wall causing atony and dilatation of bowel loops.
Osteomyelitis	An infection of bone, usually bacterial.
Priapism	Involuntary, persistent erection of the penis.

Spinal cord injuries are relatively uncommon and predominantly affect young males (80–85%). Overall, about 55% of injuries affect the cervical spine and 15% each affect the thoracic, thoracolumbar, and lumbosacral spine.

Apart from weakness, paralysis, and sensory changes, in the acute stage (depending on the level of injury) certain physiological changes are often evident. These include spinal shock, paradoxical respiration,

low body and high skin temperature, priapism, and a sweating level. Spinal shock is usually seen in injuries above the T6 level and is characterized by hypotension and bradycardia without hypovolemia due to loss of vascular sympathetic tone. The higher the level of injury and the more complete the injury, the greater the intensity and duration of spinal shock. The motor and sensory functions usually recover (if any recovery is possible) within hours, but the autonomic component of spinal shock (loss of sympathetic function) may persist for days to months. Paradoxical respiration is the result of diaphragmatic contraction in the face of intercostal muscle paralysis. The skin is warm because of a loss of sympathetically mediated peripheral vasoconstriction; this also causes the core temperature to fall. Priapism may result from an unopposed parasympathetic function, and sweating is lost because of loss of sympathetic activity.

Cardiovascular Complications

Cardiovascular complications are most common with complete injuries of the spinal cord in the cervical or high thoracic regions.

Hypovolemia and Hypotension

Spinal shock is a diagnosis of exclusion because spinal cord injury is often accompanied by other injuries. Extra emphasis should be placed on detection of any other injuries, as the lack of sensation and loss of abdominal muscle tone may make physical examination unreliable. Injuries to the chest, abdomen, pelvis, and long bones must be actively sought and excluded as a source of blood loss, hypovolemia, and hypotension. Hypovolemia and hypotension should be corrected rapidly. Hypotension in particular may increase the area of secondary injury in the spinal cord and worsen outcome. Crystalloids, colloids, and blood products may be used to replace volume as needed. Clinicians must be aware, however, that relatively low systolic blood pressure is normal in an acute high spinal cord injury, and the presence of otherwise normal cardiovascular function as indicated by adequate urinary output suggests that the patient is hemodynamically stable. Later in the course of the injury, hypotension may occur with upright posture. This is due to venous pooling and the lack of sympathetic output. It can be prevented by elastic stockings, abdominal binders, or small doses of vasopressors.

Bradycardia

Unopposed vagal activity with the loss of sympathetic activity causes bradycardia. It may be severe, especially in young patients with a high vagal tone. Atropine or glycopyrolate may be used to reverse bradycardia. If the patient is intubated, suctioning the endotracheal tube may sometimes provoke severe bradycardia.

Autonomic Hyperreflexia

Autonomic hyperreflexia is characterized by massive sympathetic outflow in response to certain noxious stimuli, usually genitourinary manipulations such as a change or irrigation of an indwelling Foley catheter. In severe cases, marked elevations in blood pressure, bradycardia, flushing, headache, nausea, and cardiovascular collapse may occur. There is severe vasoconstriction below the level of the cord injury and vasodilatation above it. This unopposed reflex sympathetic activity is usually seen well after the period of the spinal shock and is treated by withdrawing the initiating stimulus and administering vasodilators. Spinal anesthesia eliminates this reflex sympathetic outflow and should be considered for patients undergoing all major genitourinary procedures despite the patient's anatomic anesthesia below the injured level.

Venous Thrombosis

Immobilization and venous stasis, coagulation abnormalities, decreased fibrinolytic activity, and endothelial damage have been implicated in the high incidence of deep venous thrombosis in patients with complete spinal cord injury. Deep venous thrombosis usually occurs in the first 2 to 3 weeks after a spinal cord injury. A high index of suspicion is necessary, and prophylactic caval filters may be considered for high-risk patients.

Over the long term, cardiovascular function declines at an accelerated rate, and asymptomatic cardiovascular disease is also more prevalent in patients with spinal cord injury. This has been linked to the presence of dyslipidemia, enforced rest with limited exercise options, insulin resistance and hyperinsulinemia, and elevated percentages of body fat.

Respiratory Complications

Complete lesions above the C3 level are accompanied by a loss of diaphragmatic and intercostal muscle activity. These patients are dependent on artificial ventilation, initially through an endotracheal tube and later through a tracheostomy. Patients with midcervical lesions may require ventilatory assistance initially but can usually be ventilator-free long term. They remain at high risk for repeated pulmonary infections, requiring mechanical ventilation, however, due to lack of respiratory reserve. Patients with low cervical and high thoracic injuries have normal diaphragm

function but lack intercostals and abdominal muscle tone. Respiratory mechanics and efficiency are altered, as is the ability to breathe deeply or cough. Bronchial secretions are also increased because of the inability to clear airways; atelectasis, hypoxemia, and repeated infections are common.

Gastrointestinal Complications

In the acute stage of cervical or high thoracic spinal cord injury, gastric dilatation and ileus may be seen. These may persist for up to a week in complete cervical injuries. Abdominal distention, absence of bowel sounds, and abdominal x-rays showing dilated jejunal, ileal, and large bowel loops help make the diagnosis. Large volumes of fluids and electrolytes may be sequestered in the bowel and may need replacement. Cushing ulcers may also be seen in the gastrointestinal (GI) tract and may cause GI bleeding and bowel perforation. All spinal cord injury patients should receive prophylaxis with H2 blockers or proton pump inhibitors. Long-term effects on the GI tract include gastro-esophageal reflux, delayed gastric emptying, altered colonic motility, and loss of control of defecation.

Urologic Complications

Initially after high spinal cord injury, during the period of spinal shock, the bladder is completely areflexic, but the sphincter tone may be high, necessitating a Foley catheter to allow for urinary drainage. After the period of spinal shock, detrusor contractions may return, but often these movements are not coordinated with the sphincter mechanisms, resulting in poor bladder emptying. Different patterns of incoordination may be seen with conus medullaris/cauda equina injuries. Urinary infections, especially with ureteral reflux, are major problems causing renal damage, and management is targeted to the avoidance of these. Clean intermittent catheterization remains the cornerstone of management.

Skin Complications

Pressure ulcerations are an area of breakdown leading to ischemia and necrosis of the skin. Prolonged exposure to pressures just slightly above capillary filling pressure initiates a series of events leading to tissue necrosis and ulceration and is common in dependent areas such as the back and ischial tuberosities. Patients with spinal cord injury are also subject to venous stasis, atrophy of fat and muscle, moist perineum from fecal and urinary incontinence, and increased shearing stresses, which increase the risk of pressure ulceration. If left untreated there is invasion of deep fascia, muscles, bone, and joints. When this occurs, osteomyelitis is common and is a source of septicemia.

Bone and Joint Complications

Removal of weight bearing as a result of spinal cord injury results in skeletal adaptive responses characterized by loss of bone mineral and a negative calcium balance. This may be prevented by exercise therapy, electrical stimulation, and antiresorptive agents. An exercise regimen will also prevent joint contractures and maintain joint mobility.

See Also the Following Articles

Hypotension, Hypovolemia, and Septic Shock; Spinal Cord Injury, Psychological Stress of.

Further Reading

Broah, S. L. (ed.) (2005). *Managing spinal cord injury: a guide to living well with spinal cord injury.* Washington, D.C.: NRH Press.

Fehlings, M. G. and Baptiste, D. C. (2005). Current status of clinical trials for acute spinal cord injury. *Injury* **36** (supplement 2), B113–B122.

Heimburger, R. F. (2005). Return of function after spinal cord transaction. *Spinal Cord* **43**, 438–440.

Krause, J. S., Winkler, T., Steins, S. A., et al. (2001). *Spinal cord injury desk reference: guidelines for life care planning and case management.* New York: Demos Medical Publishing.

Liverman, C. T., Altevogt, B. M., Joy, J. E., et al. (eds.) (2005). *Spinal cord injury: progress, promise and priorities.* Washington, D.C.: National Academies Press.

Norenberg, M. D., Smith, J. and Marcillo, A. (2004). The pathology of human spinal cord injury: defining the problems. *Journal of Neurotrauma* **21**, 429–440.

Spinal Cord Injury, Psychological Stress of

S D Schnakenberg-Ott and M R Lovell
University of Pittsburgh Medical Center, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

Prevalence
Psychological Effects of Spinal Cord Injury
Psychological Assessment of Patients with Spinal Cord Injury
Treatment of Emotional Distress in Spinal Cord Injury Patients
Life Satisfaction and Coping with Injury

Glossary

Anti-depressants	The pharmacological classification of drugs typically used to treat mood disorders; include serotonin reuptake inhibitors and tricyclics.
Cognitive-behavioral therapy	An approach to psychotherapy that employs a variety of techniques in an attempt to change behavior or mood by disputing negative or irrational cognitive thought processes and replacing them with more rational or positive thought processes.
Neuropsychological evaluation	An assessment that typically involves the administration of several standardized cognitive tests that are sensitive to the effects of brain dysfunction.
Paralysis	A complete or partial loss of function.
Paraplegia paralysis	The paralysis of the lower extremities as the result of an injury below the first thoracic vertebrae.
Posttraumatic stress disorder (PTSD)	An anxiety disorder that is characterized by the reexperiencing of an extremely traumatic event such as a natural disaster, war, or rape. This syndrome is often accompanied by increased arousal, recurrent thoughts or images, and avoidance of stimuli associated with the trauma.
Premorbid personality characteristics	The level of psychological or personality functioning before the onset of a disorder or condition.
Quadriplegia	The paralysis of all four extremities as the result of an injury above the thoracic vertebrae.
Spinal cord injury (SCI)	Trauma to the spinal cord from a sudden blow to the spine resulting in vertebral fractures and dislocations; can result in varying degrees of paraplegia and quadriplegia.

Spinal cord injury (SCI) is generally due to a sudden blow to the spine resulting in vertebral fractures and dislocations. This type of trauma can result in varying degrees of paraplegia (the paralysis of the lower extremities as the result of an injury below the first thoracic vertebrae) and quadriplegia (the paralysis of all four extremities caused by an injury above the thoracic vertebrae). Consequently, SCI victims may endure problems with breathing and bowel or bladder control. Vulnerability to secondary medical problems such as chronic pain, respiratory problems, heart failure, and decubitus ulcers are also very common among SCI patients.

Prevalence

There is an estimated 11 000 new cases of SCI reported each year in the United States, and in 2005, there were reportedly 250 000 Americans living with SCI. Motor vehicle accidents account for a majority of the injuries, followed by falls, acts of violence (e.g., gunshot wounds and assaults), and sport activities.

Young adult men are primarily affected by SCIs; however, the rates are increasing among women. A 1998 study conducted by Schackelford, Farley, and Vines comparing men and women with SCI found that approximately 82% of SCIs occur in men compared to a 20% occurrence rate in women. The prevalence of SCIs also appears higher in Whites (62.9%), followed by African Americans (22%) and Hispanics (12.6%). The life expectancy rate for SCI patients has steadily increased during the past 50 years, primarily as a result of advanced medical care and the presence of emergency rehabilitation services. Approximately 85% of SCI patients go on to live 10 years or more if they survive the first 24 h after their injury. Recent statistics indicate that the average age of people living with SCI is 40 years. Because SCI is a lifelong disability, the need for heath-care services will continue to grow as patients with SCI live longer.

Psychological Effects of Spinal Cord Injury

In addition to secondary medical complications, many SCI patients also experience emotional problems such as depression, anxiety, and social isolation. Psychological distress may stem from diminished independence, functional impairment, chronic pain, and loss of social and work status. Depression is thought to be the most common psychological disorder among SCI

patients. A review of the literature suggests that depression rates among SCI patients vary, with a majority of the reports indicating that between 20 and 45% of people with SCI suffer from depression. Individuals with a history of psychological disorders and substance use may be at increased risk of experiencing depression following SCI.

People with SCI appear to be at greater risk for developing anxiety disorders compared to the general population. Scivoletto et al. found in 1997 that SCI patients who presented with greater medical needs were often more anxious and depressed than SCI patients with less severe medical problems. High levels of psychological distress were also indicated for SCI patients who lived alone, had lower levels of education, and were unemployed. Individuals with a disability (e.g., polio or SCI) were also found to report lower levels of life satisfaction compared to nondisabled people. In a 1999 study examining the correlates of stress in a sample of 187 SCI patients, Gerhart and colleagues found that depressive symptoms, poor perceived well-being and poor life satisfaction were strongly associated with earlier and future stress. Interestingly, medical conditions such as pressure sores and upper-extremity pain were not found to be associated with stress in this sample.

Because SCI is often the result of a traumatic and life-threatening accident, patients with SCI may also suffer from posttraumatic stress disorder (PTSD). Radnitz et al. reported a PTSD prevalence rate of 10–15% among SCI patients, whereas other researchers reported rates up to 40%. Interestingly, Radnitz and colleagues concluded that individuals with quadriplegia were diagnosed with PTSD less frequently than individuals with paraplegia, therefore suggesting that there may be an increased vulnerability for PTSD among paraplegic SCI patients. According to the researchers, the aforementioned finding may reflect the fact that quadriplegia involves injury to the upper region of the spinal cord, thereby damaging the nerve fibers that affect sympathetic arousal thought to be associated with PTSD. In a 2003 study that attempted to identify the risk factors of PTSD in SCI patients, Nielsen did not find significant relationships between level of paralysis and PTSD or trauma etiology and PTSD. However, when considering demographic variables, the author found that marriage was associated with a decreased likelihood of SCI patients developing PTSD.

There is a high risk of pre-injury substance abuse among SCI patients compared to the general population. Many patients undergo some sort of substance abuse counseling as part of a postinjury rehabilitation program; however, this does not always prevent patients from returning to substance use following hospital discharge. In addition, one 2004 study found that substance use at the time of injury was related to a negative response to problem solving in SCI patients. SCI patients are often prescribed narcotic and analgesic medication to treat chronic pain, and unfortunately, the potential for abuse of prescription drugs by some SCI patients is high. In a study examining alcohol and marijuana abuse in a sample of 123 SCI patients, it was found that patients who abused alcohol perceived their health status to be poorer, were more clinically depressed, and experienced more stress in their lives than patients with no alcohol abuse. Marijuana users also reported higher levels of stress and depression.

Suicide rates among SCI patients are five times higher than among nondisabled people, and suicide is the leading cause of death in the SCI population. Pre-injury personality characteristics, inadequate coping skills, history of substance abuse, and a poor social network have been identified in the literature as risk factors for suicidal ideation in patients with SCI. Young, adult, White males appear to have the highest rate of suicide within the SCI population. It is not surprising that most SCI patients exhibiting suicidality meet diagnostic criteria for depressive disorders. Kishi et al. examined acute-onset suicidal ideation and delayed-onset suicidal ideation in a group of SCI patients as well as patients suffering from stroke, brain injury, or myocardial infarction. Compared to patients who were identified as having delayed-onset suicidal ideation, patients with acute-onset suicidal ideation exhibited greater risk factors, including a history of psychological problems and substance abuse as well as poor social support. Therefore, it is possible that suicide rates could be significantly reduced if patients at risk are identified early in rehabilitation programs.

Psychological Assessment of Patients with Spinal Cord Injury

A comprehensive evaluation, including a thorough assessment of personality, substance use, and emotional and cognitive functioning, should be conducted for patients with acute SCI. With respect to personality testing, it may be important to identify characteristics that may interfere with rehabilitation and those individuals at risk for developing significant emotional difficulties following SCI. Chronic pain has been associated with depression and negatively affects coping. Therefore, assessing pain in SCI patients should also be an important component to the postinjury psychological evaluation. Pain could also negatively impact the rehabilitation process, especially if uncontrolled or untreated. The presence of chronic pain can also lead some patients to self-medicate with alcohol or illicit drugs. Because people

with SCI are often prescribed narcotics for pain relief, the potential of abusing prescription drugs is often present, thereby requiring physicians and mental health providers to routinely assess for improper use of prescription drugs.

A majority of SCI patients often have an underlying head injury and demonstrate poor performance on neuropsychological tests. A brain injury can be associated with lingering neurocognitive and neurobehavioral sequelae that can negatively impact rehabilitation. Therefore, following SCI, patients may need to undergo at least a brief assessment of cognitive functioning. Patients exhibiting significant cognitive difficulties on brief assessment or patients with any known brain injury should be referred for a more comprehensive cognitive evaluation, preferably by a neuropsychologist. A neuropsychological evaluation may provide additional information regarding the patient's cognitive strengths and weaknesses that can aid in determining prognosis for rehabilitation. Schmitt and Elliott found that SCI patients with higher verbal learning skills reported greater levels of self-adjustment and acceptance of their injury. A neurocognitive evaluation may also aid in determining what academic or work accommodations need to be implemented should the patient be able to resume school instruction or employment.

Treatment of Emotional Distress in Spinal Cord Injury Patients

Psychological services should be available for injured individuals and their families while the patients are in the hospital and following discharge. Psychotherapeutic intervention should focus on adjustment, self management, and skills training. Several studies conducted by Craig and colleagues have shown that employing cognitive-behavioral strategies as a form of treatment for SCI patients with significant levels of depression and anxiety are effective. Cognitive-behavioral therapy attempts to change behavior or mood by disputing negative or irrational cognitive thought processes and replacing cognitive distortions with more rational or positive thought processes. Cognitive-behavioral techniques are not only beneficial in the treatment of mood disorders, but have also been suggested efficacious for managing pain. Overall, cognitive-behavioral therapy may be potentially cost effective by decreasing the number of hospital readmissions and drug use.

In addition to individual psychotherapeutic intervention, group therapy may also be a valuable form of treatment for SCI patients, especially in a hospital setting. Groups that are psychoeducational in nature and focus on specific themes, such as coping with the injury, vocational issues, sexual dysfunction, substance abuse, or secondary medical complications may be of great benefit to acutely injured patients. Therapists might also consider inviting family members to attend a separate support group to address the emotional, logistical, and financial issues associated with having a family member with SCI.

Issues related to sexual functioning in the SCI population is becoming more recognized and discussed throughout SCI literature. Of course, the level of physical impairment or paralysis following SCI can affect sexual functioning. Psychotherapeutic intervention that addresses both adaptation and coping with the physical and functional changes in sexuality following SCI is necessary. Erectile dysfunctions in men and physiological changes in women are common features of SCI. Obviously, people with SCI are not immune to or incapable of acquiring sexually transmitted diseases (STDs), and women with SCI may become pregnant. Therefore, mental health providers may also find themselves in a position to educate SCI patients on safe sex practices.

Pharmacological treatment may also be advantageous for the treatment of emotional disorders (e.g., depression and anxiety) in patients with SCI. Because they have fewer side effects compared to tricyclic antidepressants, serotonin reuptake inhibitors (SSRIs) usually have a more benign side-effect profile. However, few studies have been conducted that demonstrate the efficacy and possible side effects of antidepressant therapy in SCI patients.

Life Satisfaction and Coping with Injury

Factors such as premorbid personality characteristics and available social support can strongly influence coping with SCI. Frank and Elliott found that patients who believed they had some control over their situation and events experienced less depression and were better adjusted than patients who believed they had little control over their life situation and health status. Other protective factors against depression and anxiety appear to be age and the number of years since the injury occurred.

Life satisfaction has been shown to be higher in SCI patients with adequate social support than in SCI patients with poor social support systems. Length of time living with SCI was also positively correlated with greater life satisfaction. Similarly, a 2002 study found that social support was associated with lower levels of hopelessness and depression. Ongoing health-care services are also necessary because the life span of individuals with SCI continues to rise, thereby increasing the probability of secondary medical conditions. One study found that SCI patients who received regular follow-up medical care reported better health, independence, and fewer

severe secondary medical complications, and demonstrated lower depression scores than SCI patients who did not receive follow-up care. In essence, adequate health care delivered by well-trained physicians, nurses, mental heath-care providers, vocational rehabilitation specialists, and other essential staff is pertinent in order to help facilitate a longer life expectancy and increased quality of life among SCI survivors.

See Also the Following Articles

Depression Models; Spinal Cord Injury, Physical Stress of.

Further Reading

Beedle, A. and Kennedy, P. (2002). Quality of social support predicts hopelessness and depression post spinal cord injury. *Journal of Clinical Psychology in Medical Settings* **9**(3), 227–324.

Craig, A. and Hancock, K. (1998). Living with spinal cord injury: longitudinal factors, interventions, and outcomes. *Clinical Psychology and Psychotherapy* **5**, 102–108.

Craig, A., Hancock, K. and Dickson, H. (1994). Spinal cord injury: a search for determinants of depression 2 years after the event. *British Journal of Clinical Psychology* **33**, 221–230.

DeVivo, M., Black, K., Richards, S., et al. (1991). Suicide following spinal cord injury. *Paraplegia* **29**, 620–627.

Dijkers, M., Abela, M. and Gordon, W. (1995). The aftermath of spinal cord injury. In: Stover, S. L., DeLisa, J. A. & Whiteneck, G. G. (eds.) *Spinal cord injury: clinical outcomes from the model systems*, pp. 185–212. Gaithersburgh MD: Aspen.

Dreer, L., Elliott, T. and Tucker, E. (2004). Social problem-solving abilities and health behaviors among persons with recent onset spinal cord injury. *Journal of Clinical Psychology in Medical Settings* **11**(1), 7–13.

Dunn, M., Love, L. and Ravesloot, C. (2000). Subjective health in spinal cord injury after outpatient healthcare follow-up. *Spinal Cord* **38**, 84–91.

Frank, R. and Elliott, T. (1989). Spinal cord injury and health locus of control beliefs. *Paraplegia* **27**, 250–256.

Gerhart, K., Weitzenkamp, D., Kennedy, P., et al. (1999). Correlates of stress in long term spinal cord injury. *Spinal Cord* **37**, 183–190.

Hess, D., Marwitz, J. and Kreutzer, J. (2003). Neuropsychological impairments after spinal cord injury: a comparative study with mild traumatic brain injury. *Rehabilitation Psychology* **48**, 151–156.

Kemp, B. and Krause, J. (1999). Depression and life satisfaction among people ageing with post-polio and spinal cord injury. *Disability and Rehabilitation* **21**, 241–249.

Kennedy, P. (2001). Spinal cord injuries. In: Johnston, D., Johnston, M., Bellack, A. & Hersen, M. (eds.) *Health psychology. Vol. 8: Comprehensive clinical psychology*, pp. 445–462. Amsterdams: Elsevier Science.

Kennedy, P., Lowe, R., Grey, N., et al. (1995). Traumatic spinal cord injury and psychological impact: across sectional analysis of coping strategies. *British Journal of Clinical Psychology* **34**, 627–639.

Kishi, Y., Robinson, R. and Kosier, J. (2001). Suicidal ideation among patients during the rehabilitation period after life threatening physical illness. *Journal of Nervous and Mental Disease* **189**, 623–628.

McColl, M. A. and Rosenthal, C. (1994). A model of resource needs of aging spinal cord injured men. *Paraplegia* **32**, 261–270.

National Spinal Cord Injury Statistical Center (1995). *Spinal cord injury: facts and figures at a glance*. Birmingham, AL: National Spinal Cord Injury Statistical Center.

Nielsen, M. S. (2003). Prevalence of posttraumatic stress disorder in persons with spinal cord injuries: the mediating effect of social support. *Rehabilitation Psychology* **48**, 289–295.

North, N. (1999). The psychological effects of spinal cord injury: a review. *Spinal Cord* **37**, 671–679.

Radnitz, C., Hsu, L., Tirch, D., et al. (1998). A comparison of posttraumatic stress disorder in veterans with and without spinal cord injury. *Journal of Abnormal Psychology* **107**, 676–680.

Richards, J. S. (2005). Spinal cord injury pain: impact, classification, treatment trends, and implications from translational research. *Rehabilitation Psychology* **50**, 99–102.

Richards, J. S., Kewman, D. G. and Pierce, C. A. (2000). Spinal cord injury. In: Frank, R. G. & Elliott, T. R. (eds.) *Handbook of rehabilitation psychology*, pp. 11–27. Washington DC: American Psychological Association.

Schmitt, M. M. and Elliott, T. R. (2004). Verbal learning ability and adjustment to recent-onset spinal cord injury. *Rehabilitation Psychology* **49**, 288–294.

Scivoletto, G., Petrelli, A., Di Lucente, L., et al. (1997). Psychological investigation of spinal cord injury. *Spinal Cord* **35**, 516–520.

Shackelford, M., Farley, T. and Vines, C. L. (1998). A comparison of women and men with spinal cord injury. *Spinal Cord* **36**, 337–339.

Young, M. E., Rintala, D. H., Rossi, D., et al. (1995). Alcohol and marijuana use in a community-based sample of persons with spinal cord injury. *Archives of Physical Medicine and Rehabilitation* **76**, 525–532.

Relevant Website

Neurosurgery Today (2003). Spinal cord injury FAQ. http://www.neurosurgerytoday.org.

Spousal Abuse *See:* Domestic Violence; Marital Conflict; Marital Status and Health Problems.

Startle Response

C O Ladd, P M Plotsky and M Davis
Emory University Medical School, Atlanta, GA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 486–494, © 2000, Elsevier Inc.

Characterization of the Startle Response
Mediation of the Startle Reflex
Modulation of the Startle Reflex By Stress
Other Startle Phenomena

Glossary

Amygdala	A bilateral limbic structure located just anterior to the ventral tip of the hippocampus. The amygdala is a collection of perhaps 13 different nuclei, including the central nucleus of the amygdala, the basolateral amygdala nucleus, and the lateral amygdala nucleus. This structure is critically involved in fear as well as attention, aggression, sexual behavior, and certain aspects of appetitive behavior, such as secondary reinforcement.
Bed nucleus of the stria terminalis (BNST)	A structure surrounding the anterior commissure that is related anatomically and functionally to the amygdala. It is also believed to play a role in fear or anxiety as well as sexual arousal.
Corticotropin releasing hormone (CRH)	A 41-amino-acid peptide that coordinates the behavioral, endocrine, and autonomic indices of the mammalian stress response by acting as a neurohormone and neurotransmitter (neuromodulator) in the brain. CRH and its receptors are expressed in both the amygdaloid nuclei and the BNST.
Nucleus reticularis caudalis pontis	A group of large neurons in the pontine reticular formation that forms the final nucleus in the brain stem involved in the primary acoustic startle circuit.

The mammalian startle response consists of a rapid extension and then flexion of a series of muscles in response to an abrupt stimulus that is usually auditory but may be somatosensory, vestibular, or visual. As a simple, stimulus-controlled behavior, the mammalian startle response has been used to investigate the neuroanatomical and cellular basis for learning and behavior.

Characterization of the Startle Response

Acoustic Startle

The acoustic startle reflex (ASR) typically is elicited by a brief auditory stimulus (90–120 dB, 50 ms duration). The minimum stimulus required to elicit a response in normal rats is approximately 80–90 dB, provided that the stimulus reaches full intensity within approximately 12 ms of the stimulus onset. Even very intense auditory stimuli do not elicit a startle reflex if they have long rise times (e.g., >25 ms). In mammals, the amplitude of the startle response is proportional to the duration of the stimulus (up to 8 ms) with a direct positive relationship between startle amplitude and stimulus intensity. In rodents, high-frequency auditory stimuli are more effective than low-frequency stimuli in eliciting startle. The startle reflex has been observed in every mammal tested and is characterized by an extremely short latency. In rats, electromyographic (EMG) recordings in hindlimb muscles indicate a response latency of 8 ms following onset of an auditory stimulus.

Air Puff Startle

The startle reflex can also be elicited by an air puff that contains both tactile and acoustic components. Across laboratories, the duration of the air puff stimulus ranges from 100 to 120 ms and the stimulus intensity ranges from 9 to 75 psi. Studies in hearing-impaired rodents have shown that the auditory component of a low-intensity air puff stimuli (12.5 psi) is primarily responsible for eliciting the startle response. With more intense air puff stimuli, the auditory and tactile stimuli synergize to produce large-amplitude startle responses. Air puff startle has been used not only to elicit and measure the startle reflex, but also as an unconditioned stressor. For example, in rats, air puff stimuli can induce ultrasonic vocalizations, activate the hypothalamic-pituitary-adrenal (HPA) axis, and increase defensive withdrawal responses.

Human Startle

In humans, startle involves primarily a flexion response that is most pronounced in the upper half of the body, particularly the face, neck, and shoulders, with the minimum response being flexion of the sternocleidomastoid muscle (a latency of 14 ms). The response begins with blinking of the eyes (a latency of 35 ms) and extends caudally, causing forward head movement, facial grimacing with widening of the mouth and occasional baring of the teeth, elevation

of the shoulders, abduction of the arms, bending of the elbows, pronation of the forearms, flexion of the fingers and trunk, abdominal contraction, and knee flexion. Thus, the human startle response primarily involves flexor muscles, but may also include extensors, depending on the subject's posture at the time of stimulus onset. Startle in humans is almost always measured via the amplitude of the eyeblink, which is highly graded in magnitude and shows the same types of plasticity as the whole body response in rodents.

Rodent Startle

In rodents, startle can result in prominent extension of the forelimbs and hindlimbs to the point where all four feet come off the ground. Smaller startle responses involve flexion of the back so that the body shortens and rises into a characteristic hunched posture. In both cases, startle amplitude is measured as a downward movement of a cage that is detected by transducers sensitive to rapid movement that put out a voltage proportional to the magnitude of the startle reflex (e.g., accelerometers, piezoelectric disks, load cells, and phonograph cartridges). Startle amplitude is typically defined as the maximal transducer output during a brief period of time (e.g., 200 ms) after onset of the eliciting stimulus. By sampling cage output over brief periods of time, startle can be readily discriminated from spontaneous or drug-induced changes in locomotor or stereotyped behavior.

Mediation of the Startle Reflex

Anatomical Structures

Because the ASR has such a short latency, a simple neural pathway must mediate it. In 1982, it was proposed that acoustic startle in the rat was mediated by four synapses: three in the brain stem (the ventral cochlear nucleus, an area just medial and ventral to the ventral nucleus of the lateral lemniscus, and the nucleus reticularis pontis caudalis) and one on to motoneurons in the spinal cord. Electrolytic lesions of these nuclei eliminated acoustic startle, whereas single-pulse electrical stimulation of these nuclei elicited startlelike responses with a progressively shorter latency as the electrode was moved farther down the startle pathway. However, it has now been found that the selective destruction of cell bodies without a concomitant loss of fibers of passage in the ventral nucleus of the lateral lemniscus or the area just ventral and medial to it did not eliminate startle, whereas comparable lesions in the nucleus reticularis pontis caudalis still completely eliminated startle. These data strongly suggested, therefore, that an additional synapse in the ventral nucleus of the lateral lemniscus or the area just ventral and medial might not be necessary for relaying auditory information from the cochlear nucleus to the pontine reticular formation. However, all data still pointed to the critical importance of the nucleus reticularis pontis caudalis in the acoustic startle reflex. Until very recently, it has been unclear how auditory information gets to this traditionally nonauditory part of the brain stem. It is now known that in both rodents and humans there is a small group of very large cells (35 μm in diameter) embedded in the cochlear nerve, called cochlear root neurons. These neurons receive direct input from the spiral ganglion cells in the cochlea, making them the first acoustic neurons in the central nervous system. They send exceedingly thick axons (sometimes as wide as 7 μm) through the trapezoid body, at the very base of the brain, directly to an area just medial and ventral to the lateral lemniscus and continue on up to the deep layers of the superior colliculus. However, they give off thick axon collaterals that terminate directly in the nucleus reticularis pontis caudalis, exactly at the level known to be critical for the acoustic startle reflex. Most important, the chemical destruction of these cochlear root neurons completely eliminates the acoustic startle reflex. Although damage to the auditory root, where the cochlear root neurons reside, has not been fully ruled out, other tests indicated that these animals could clearly orient to auditory stimuli (e.g., suppression of licking) and had normal compound action potentials recorded from the cochlear nucleus.

Although there is still some disagreement, the acoustic startle pathway may consist of only three synapses on to the (1) cochlear root neurons, (2) neurons in the nucleus reticularis pontis caudalis, and (3) motoneurons in the facial motor nucleus (pinna reflex and perhaps the eyeblink) or spinal cord (whole-body startle; **Figure 1**).

In humans, the obicularis oculi muscles that are innervated by the facial motor nucleus mediate the eyeblink component of startle. Because the facial motor nucleus receives a direct projection from the nucleus reticularis pontis caudalis, the eyeblink component of startle could involve a pathway consisting of cochlear root neurons, the nucleus reticularis pontis caudalis, and the facial motor nucleus. However, this simple neural pathway may not be able to account for the relatively long latency of the eyeblink reflex (35 ms), so other pathways have been suggested.

Neurotransmitters that Mediate the Startle Reflex

Auditory nerve to cochlear root neuron synapse At the present time, there are no direct data on the identity of the neurotransmitter that may mediate startle at this level. Preliminary experiments in rodents suggest it may be glutamate acting on non-*N*-methyl-D-aspartate (NMDA) receptors because a local infusion

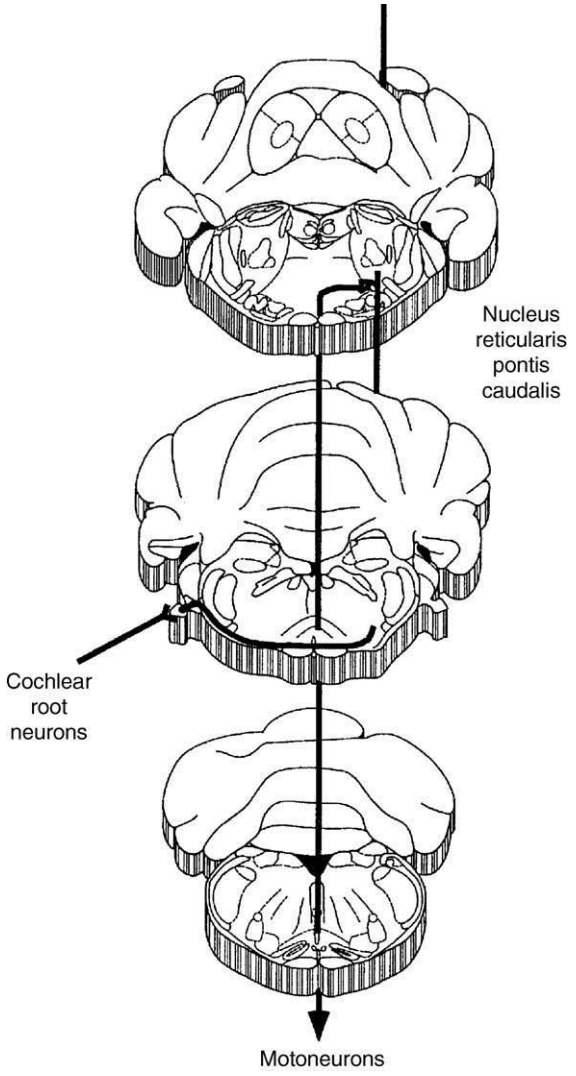


Figure 1 Schematic diagram of a primary acoustic startle pathway. Auditory nerve fibers synapse onto cochlear root neurons embedded in the auditory nerve. Axons from these cells project through the ventral acoustic stria and send projections to the nucleus reticularis pontis caudalis (PnC). Projections of cells in the PnC form the reticulospinal tract, which make monosynaptic and polysynaptic synapses in the spinal cord. These cells also project to the facial motor nucleus in areas critical for the pinna component of startle in rats and perhaps the eyeblink component of startle in humans (not shown). Lesions of the cochlear root neurons, the ventral acoustic stria, the PnC, or reticulospinal tract eliminate the acoustic startle response.

of the non-NMDA antagonist CNQX in the vicinity of the root neurons depressed startle, whereas local infusion of the NMDA antagonist AP5 did not.

Cochlear root neuron to reticulospinal neuron synapse Glutamate acting on both NMDA and non-NMDA receptors may well be the neurotransmitter that mediates startle at the level of the nucleus reticularis pontis caudalis. Local infusion of either CNQX or AP5 in rats chronically implanted with bilateral cannulas aimed at the nucleus reticularis pontis caudalis significantly reduced startle amplitude by as much as 70–80%. AP5 and CNQX attenuated startle over a dose range, which indicated that the nucleus reticularis pontis caudalis is extremely sensitive to these compounds. Thus, small changes in neurotransmitter activity in this area may have a profound effect on the expression of the acoustic startle reflex.

Reticulospinal neurons to motor neuron synapse Intrathecal infusion (into the subarachnoid space of the spinal cord to allow drugs to diffuse in the vicinity of motoneurons) of the non-NMDA antagonist CNQX or the NMDA antagonist AP5 each reduced the amplitude of the whole-body startle reflex in a dose-dependent manner. When EMG activity in the hindlimbs was used to define startle, intrathecal administration of a combination of AP5 and CNQX completely eliminated all EMG components of the startle response, consistent with the hypothesis that excitatory amino acids actually mediate startle at the spinal level. Moreover, when the EMG components of the startle reflex were separated into early or late components, CNQX preferentially eliminated the early components (a mean latency of 8.29 ms), whereas AP5 preferentially eliminated the later components (a mean latency of 14.57 ms or longer).

The fact that startle appears to involve both non-NMDA and NMDA receptors at the spinal level can be interpreted in a number of ways. The simplest possibility is that glutamate is released from reticulospinal tract terminals, which form monosynaptic connections onto motoneurons expressing both NMDA and non-NMDA receptors. The short-latency EMG components then reflect the activation of non-NMDA receptors, whereas the longer-latency EMG components reflect the activation of NMDA receptors. A second possibility is that reticulospinal neurons make both monosynaptic and polysynaptic connections via interneurons onto spinal motoneurons. The short-latency EMG components then reflect the monosynaptic connections on to non-NMDA receptors, whereas the longer-latency EMG components reflect polysynaptic connections involving NMDA receptors. A third possibility is that the short-latency EMG components of startle reflect the activation of non-NMDA receptors by the short-latency startle pathway described previously, whereas the later components reflect activity in other brain-stem structures, which then eventually activate NMDA receptors on spinal motoneurons. This idea can be tested by the direct stimulation of reticulospinal neurons in conjunction with intrathecal infusion of excitatory amino acid antagonists.

In summary, the excitatory amino acid glutamate may well mediate startle at each of the three central synapses along the acoustic startle pathway. NMDA and non-NMDA receptors appear to be involved at both the spinal cord and nucleus reticularis pontis caudalis. Processes involved in the release of glutamate, as well as changes in the sensitivity of glutamate receptors, are expected to affect startle and startle plasticity.

Modulation of the Startle Reflex By Stress

The Fear-Potentiated Startle Effect

In rats, the amplitude of the startle reflex is increased when elicited by an auditory stimulus shortly after the onset of a light paired previously with a foot shock. This fear-potentiated startle effect occurs only following prior light–shock pairings and not when lights and shocks have been presented in an unpaired or random relationship, indicating its dependence on prior Pavlovian fear conditioning. It is probable that the light produces a state of fear, which increases reflexive behavior, because drugs such as buspirone or valium, which reduce fear and anxiety in humans, block the increase in startle in the presence of the light but do not systematically alter startle in the absence of light when appropriate doses are used. When the eyeblink component of the startle reflex is measured in humans, fear-potentiated startle can be produced using very similar conditioning procedures to those used in rodents. Humans also show an increase in startle amplitude simply when they are told that a shock might occur after presentation of a light or when they see threatening pictures such as a gun or an attacking dog. Conversely, pleasant pictures shown to humans or a light paired with food shown to rats reduce startle amplitude.

Figure 2 summarizes our knowledge of the neural pathways involved in the fear-potentiated startle effect using a visual conditioned stimulus. Visual information goes through the visual thalamus to the perirhinal cortex, which then projects to the amygdala, which has a direct projection to the nucleus reticularis pontis caudalis in the startle pathway. The lateral posterior nucleus of the thalamus, which receives heavy input from the retina, projects directly to the perirhinal cortex. Lesions of the lateral posterior nucleus of the thalamus, combined with lesions of the dorsal and ventral nuclei of the lateral geniculate, completely block

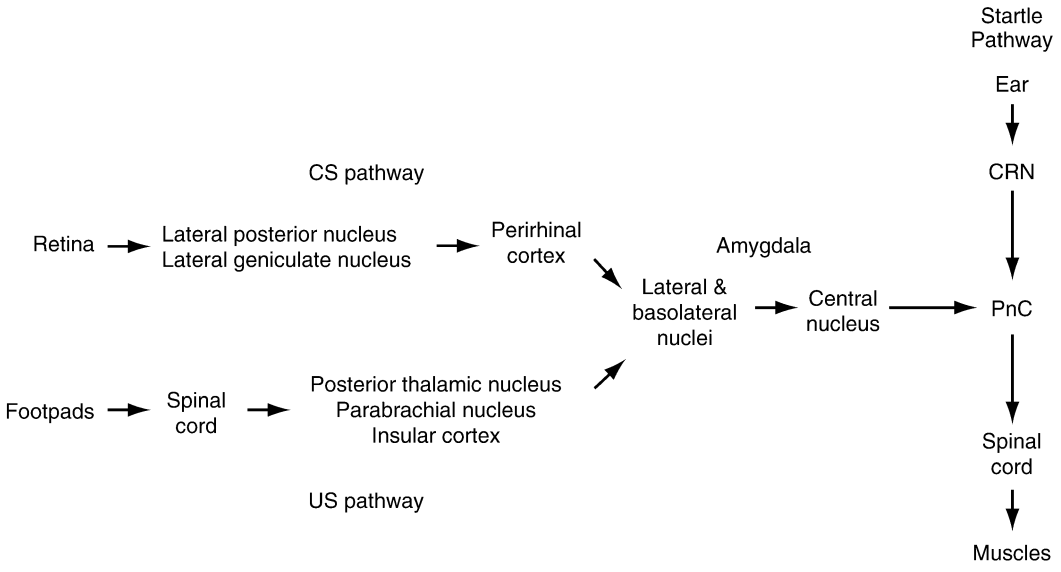


Figure 2 Working hypothesis of the neural pathways involved in fear-potentiated startle using a visual-conditioned stimulus and a foot shock-unconditioned stimulus. Visual information from the retina goes in parallel to the dorsal lateral geniculate nucleus and the lateral posterior nucleus in the thalamus. Lesions of both of these nuclei together after light–shock pairings eliminate fear-potentiated startle to a light. Lesions of either area alone or lesions of the visual cortex alone, which receive inputs from the dorsolateral geniculate nucleus, do not block the expression of fear-potentiated startle to a light. The lateral posterior nucleus projects directly to the perirhinal cortex, whereas the dorsolateral nucleus projects to the perirhinal cortex indirectly via the visual cortex. The perirhinal cortex projects to lateral and basolateral nuclei of the amygdala, which project to the central nucleus. The central nucleus has direct projections to the nucleus reticularis pontis caudalis, an obligatory part of the acoustic startle pathway; however, indirect projections via a synapse in the rostral midbrain probably are involved as well. Shock information is relayed to the lateral and basolateral amygdala via parallel inputs arising from the insular cortex and posterior thalamic nucleus. Lesions of both these areas together prior to light–shock training block the acquisition of fear-potentiated startle. However, lesions of both of these areas following training have no effect. Lesions of either area alone fail to block acquisition or expression of fear-potentiated startle. CRN, cochlear root neurons; CS, conditioned stimulus; PnC, nucleus reticularis pontis caudalis; US, unconditioned stimulus.

fear-potentiated startle using a visual-conditioned (but not an auditory-conditioned) stimulus. Lesions of either structure alone have no effect. Furthermore, posttraining lesions of the visual cortex, to which the dorsal nucleus of the lateral geniculate projects, do not affect fear-potentiated startle. Hence, visual information gets to the perirhinal cortex either directly from the lateral posterior nucleus or indirectly from the dorsal nucleus of the lateral geniculate via the visual cortex. Lesion data suggest that either one of these pathways is sufficient to mediate fear-potentiated startle using a visual-conditioned stimulus. The perirhinal cortex has heavy reciprocal projections with the lateral and/or basolateral nuclei of the amygdala. Relatively small lesions of the perirhinal cortex completely block fear-potentiated startle, provided the lesion includes an area of perirhinal cortex both dorsal and ventral to the rhinal sulcus.

Selective destruction of cell bodies in the lateral and basolateral nuclei of the amygdala completely block fear-potentiated startle using either a visual- or an auditory-conditioned stimulus when the lesions were made either before or after training. Lesions of the central nucleus (CeA), innervated by the basolateral nucleus of the amygdala, block the expression of fear-potentiated startle using either a visual- or an auditory-conditioned stimulus. The CeA has both direct and indirect projections via the central gray to the nucleus reticularis pontis caudalis, and lesions at several points along this pathway block the expression of fear-potentiated startle.

When shock is paired with light during fear-potentiated startle training, shock information is relayed to the basolateral amygdala via two parallel pathways. These include the caudal granular/dysgranular insular cortex and the posterior intralaminar nuclei of the thalamus. Pretraining lesions of both of these areas together, but not either one alone, blocked the acquisition of fear-potentiated startle. However, posttraining combined lesions of these same brain areas did not prevent the expression of conditioned fear.

Sensitization of Startle by Foot Shocks

When rats are presented with a series of startle stimuli before and after a series of 10 foot shocks presented at a 10-s interstimulus interval, startle shows a progressive increase in amplitude that peaks in approximately 10 min. The magnitude of this startle sensitization effect is directly related to the number or intensity of intervening shocks. Moreover, shocks given prior to presenting any startle stimuli also lead to an elevation of startle, indicating that enhanced startle results from sensitization rather than simply from dishabituation. Like fear-potentiated startle, lesions of the amygdala block the shock sensitization of startle.

Noise-Enhanced Startle

The startle reflex is increased when elicited in the presence of steady background noise as well as following a period of sustained high-level noise. Although this may reflect an effect of noise stress, noise-enhanced startle is not blocked by lesions of the amygdala and shows inconsistent sensitivity to anxiolytic compounds.

Light-Enhanced Startle

When the startle reflex is elicited by a loud sound 5–20 min after a bright light has been turned on, there is an increase in the startle reflex. This effect is also decreased significantly by drugs that reduce anxiety in humans and may indicate an unconditioned anxiogenic effect of bright light that enhances startle. Humans show a significant increase of startle amplitude (i.e., of the eyeblink response) in the dark. The species difference may reflect fear of the light in a nocturnal species (rats) compared to fear of the dark in a diurnal species (humans). When the lights go off, many people feel more anxious, especially if they were afraid of the dark when they were young. In patients with posttraumatic stress disorder, startle is increased in the dark to a greater degree than that seen in combat controls. These individuals report that the darkness makes them think of being back at their guard post in Vietnam and anxious about being hit by an incoming mortar.

Like fear-potentiated startle, local infusion of the glutamate antagonist NBQX into the basolateral nucleus of the amygdala blocks light-enhanced startle. However, unlike fear-potentiated startle, NBQX infused into the BNST, but not the CeA, blocks light-enhanced startle. The BNST also receives input from the basolateral nucleus of the amygdala and projects directly to the startle pathway at the level of the nucleus reticularis pontis caudalis. Thus, a pathway going from the visual thalamus to basolateral amygdala to BNST to nucleus reticularis pontis caudalis may mediate light-enhanced startle.

Corticotropin Releasing Hormone-Enhanced Startle

The intraventricular administration of the peptide CRH produces a variety of behavioral and neuroendocrine effects similar to those seen during fear and anxiety. Intraventricular administration of the CRH antagonist α-helical CRH(9–41) blocks the behavioral and neuroendocrine effects of natural stressors

or conditioned fear. The intraventricular administration of CRH increases the acoustic startle reflex. This effect can be blocked by the benzodiazepine chlordiazepoxide, suggesting that the excitatory effect of CRH on startle reflects an anxiogenic effect of the hormone. This effect still occurs after adrenalectomy, indicating that it is not dependent on the release of corticosterone from the adrenal glands. Thus, CRH-enhanced startle represents yet another example of a long-lasting, unconditioned, anxiogenic effect. Chemically induced lesions of the BNST, but not of the amygdala, block CRH-enhanced startle. Moreover, local infusion of CRH into the BNST, but not into the CeA, increases startle amplitude. Hence, CRH-enhanced startle is more similar to light-enhanced startle than fear-potentiated startle in terms of its dependence on the BNST versus the CeA.

Drug Withdrawal

Several studies in both rats and humans have shown that alcohol withdrawal increases acoustic startle amplitude. In rats, nicotine withdrawal or withdrawal from chronic diazepam also increases startle amplitude. In contrast, naloxone-precipitated opiate withdrawal depresses startle markedly. Withdrawal from stimulants such as cocaine or amphetamine has little effect on startle.

Maternal Separation

Exposing rat pups to moderate periods of maternal separation (180 min/day during postnatal days 2–14) increases the adult HPA response to acute air puff startle compared to nonhandled or normally reared rat pups. These animals also show an elevation of the acoustic startle reflex. In contrast, brief (15-min) periods of maternal separation during that same neonatal period renders these animals hyporesponsive to air puff startle (and other psychological stressors) as adults compared to nonhandled rat pups. This alteration in sensitivity and reaction to an acute air puff stimulus may be due to persistent changes in the central GABA, norepinephrine, and CRH systems, as well as in hippocampal glucocorticoid receptor expression.

Other Stressors

In rats, startle is increased during frustrative nonreward, but is decreased following cold swim stress or food, water, or sleep deprivation. In humans, the eyeblink component of startle is increased when elicited in the presence of noxious odors or heat (47–49°C) applied to the thumb. However, anticipation of giving a speech produces a pronounced decrease in startle amplitude, even though this is a very anxiogenic event for humans. Although the pattern of data is by no means clear, it appears that stressors that direct attention inward (food, water, sleep deprivation, and hypothermia) depress startle, whereas stressors that direct attention outward (threat of shock, a predator, and a harmful environment) tend to increase startle amplitude.

Other Startle Phenomena

Habituation and Dishabituation

When a startle stimulus is repeated relatively rapidly (e.g., once every 10–15 s), the amplitude of the startle reflex will generally decline over repeated test trials, a phenomenon called habituation. In general, the rate of response decrement within a test session is faster with short intervals than with long intervals between stimuli (interstimulus intervals) and with weak startle stimuli than with strong startle stimuli. In contrast, across-session habituation (e.g., measured several hours later or the next day) is generally greater following long intertrial intervals than short intertrial intervals and loud startle stimuli than weak startle stimuli, provided that a common test interval or test intensity is used to assess the level of habituation.

If, following habituation, a novel intense stimulus is presented, the startle amplitude may increase above its habituated level, a phenomenon called dishabituation. It is generally believed that dishabituation is a special case of sensitization and that habituation and sensitization represent separate processes that summate to form the final response amplitude at any given time. In fact, a loud startle stimulus may produce both habituation and sensitization, which jointly determine startle amplitude.

By eliciting the startle reflex electrically via the implantation of tiny electrodes at different levels of the startle pathway, researchers concluded that short-term habituation may result from a change in transmission at the nucleus reticularis pontis caudalis. Because lesions in a variety of brain areas fail to alter the rate of short-term habituation of the startle reflex, it may involve processes intrinsic to the startle pathway, such as receptor desensitization or the buildup of a very local inhibitory process.

However, brain areas extrinsic to the startle pathway itself may mediate long-term habituation. Leaton and Supple have shown that lesions of the cerebellar vermis, which is not part of the acoustic startle pathway, significantly attenuates the decrease in startle amplitude seen across daily test sessions, without affecting the rate of response decrement within test sessions. This blockade of long-term habituation was observed with two different stimulus intensities and

cannot be explained by ceiling or floor effects caused by cerebellar lesions.

Sensitization of startle may also involve brain areas extrinsic to the startle pathway. Borszcz and colleagues found that loud startle stimuli produce sensitization that can best be described as fear of the experimental context in which startle is measured. Reintroducing animals into this context produces a good deal of freezing, a reliable index of fear. Fear of the context elevates startle on subsequent test sessions, leading to a reduction in the amount of long-term response decrement. Treatments such as lesions of the amygdala or drugs such as diazepam, which reduce fear in many situations, facilitate the degree of long-term response decrement, presumably by blocking sensitization and hence allowing long-term habituation to be revealed.

Prepulse Inhibition

The amplitude of the startle response is reduced when the strong startle-eliciting stimulus is preceded shortly (30–500 ms) by a weaker stimulus, or prepulse. This phenomenon, termed prepulse inhibition (PPI), may be a measure of the gating or filtering of sensory information. PPI may be mediated by prepulse inputs to the inferior and superior colliculi, which activate the pedunculopontine tegmental nucleus. This structure, in turn, sends inhibitory cholinergic projections to the caudal pontine reticular nucleus.

A cortico-pallido-striato-pontine circuit connecting limbic cortical regions and subcortical dopaminergic systems to the pedunculopontine tegmental nucleus modulates PPI. These pathways are primarily glutamatergic, dopaminergic, or GABAergic. Both glutamate and dopamine have been implicated in the modulation of PPI, glutamate as a facilitator via subcortical NMDA receptors and dopamine as an inhibitor via D2 receptors in the nucleus accumbens. In the presence of excess mesolimbic dopamine, the pedunculopontine tegmental nucleus is inhibited both directly and indirectly (through the ventral pallidum) by accumbal efferents. The inhibition of the pedunculopontine tegmental nucleus would then block the prepulse from activating this structure and thus reduce its inhibitory effect.

The disruption of prepulse inhibition in rats by dopamine agonists or stress has received particular consideration as an animal model of the putative attentional impairments underlying the positive symptoms of schizophrenia thought to be associated with excessive dopamine transmission. Reduced PPI is also observed in other disorders characterized by sensorimotor and cognitive gating deficiencies, including Huntington's disease, obsessive-compulsive disorder, attention deficit disorder, and Tourette's syndrome. Schizophrenic patients are often unable to sustain voluntary attention or filter irrelevant thoughts and stimuli. Several neuroimaging studies of schizophrenics have demonstrated anatomical, metabolic, and cellular abnormalities in the cortico-striato-pallido-pontine circuitry modulating PPI, which may underlie these deficits. This model is currently being used to investigate the genetic and early environmental contributions to PPI deficiency, as well as to screen potential antipychotics.

See Also the Following Articles

Evolutionary Origins and Functions of the Stress Response; Fear; Fight-or-Flight Response.

Further Reading

Boulis, N. M., Kehne, J. H., Miserendino, M. J. D., et al. (1990). Differential blockade of early and late components of acoustic startle following intrathecal infusion of 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX) or D, L-2-amino-5-phosphonovaleric acid (AP-5). *Brain Research* **520**, 240–246.

Davis, M. (1984). The mammalian startle response. In: Eaton, R. C. (ed.) *Neural mechanisms of startle behavior*, pp. 287–351. New York: Plenum.

Davis, M. and Astrachan, D. I. (1978). Conditioned fear and startle magnitude: effects of different footshock or back-shock intensities used in training. *Journal of Experimental Psychology: Animal Behavior Processes* **4**, 95–103.

Davis, M., Falls, W. A., Campeau, S., et al. (1993). Fear-potentiated startle: a neural and pharmacological analysis. *Behavioural Brain Research* **58**, 175–198.

Davis, M., Gendelman, D. S., Tischler, M. D., et al. (1982). A primary acoustic startle circuit: lesion and stimulation studies. *Journal of Neuroscience* **6**, 791–805.

Davis, M., Parisi, T., Gendelman, D. S., et al. (1982). Habituation and sensitization of startle reflexes elicited electrically from the brainstem. *Science* **218**, 688–690.

Dawson, M. E., Schell, A. M. and Bohmelt, A. H. (1999). *Startle modification: implications for neuroscience, cognitive science, and clinical science*. Cambridge, UK: Cambridge University Press.

Dunn, A. J. and Berridge, C. W. (1990). Physiological and behavioral responses to corticotropin-releasing factor administration: is CRF a mediator of anxiety or stress responses. *Brain Research Reviews* **15**, 71–100.

Ebert, U. and Koch, M. (1992). Glutamate receptors mediate acoustic startle input to the reticular brain stem. *NeuroReport* **3**, 429–432.

Engelmann, M., Thrivikraman, K. V., Su, Y., et al. (1996). Endocrine and behavioral effects of airpuff startle in rats. *Psychoneuroendocrinology* **21**, 391–400.

Geyer, M. A., Swerdlow, N. R., Mansbach, R. S., et al. (1990). Startle response models of sensorimotor gating

and habituation deficits in schizophrenia. *Brain Research Bulletin* **25**, 485–498.
Grillon, C. and Davis, M. (1997). Fear-potentiated startle conditioning in humans: effects of explicit and contextual cue conditioning following paired versus unpaired training. *Psychophysiology* **34**, 451–458.
Grillon, C., Morgan, C. A., Davis, M., et al. (1998). Effect of darkness on acoustic startle in Vietnam veterans with post-traumatic stress disorder. *American Journal of Psychiatry* **155**, 812–817.
Grillon, C., Pellowski, M., Merikangas, K. R., et al. (1997). Darkness facilitates the acoustic startle reflex in humans. *Biological Psychiatry* **42**, 461–471.
Groves, P. M. and Thompson, R. F. (1970). Habituation: a dual process theory. *Psychological Review* **77**, 419–450.
Koch, M. and Schnitzler, H.-U. (1997). The acoustic startle response in rats: circuits mediating evocation, inhibition and potentiation. *Behavioural Brain Research* **89**, 35–49.
Landis, C. and Hunt, W. (1939). *The startle paradigm*. New York: Farrar and Rinehart.
Lang, P. J. (1995). The emotion probe. *American Psychologist* **50**, 372–385.
Lee, Y. and Davis, M. (1997). Role of the hippocampus, bed nucleus of the stria terminalis and amygdala in the excitatory effect of corticotropin releasing (CRH) hormone on the acoustic startle reflex. *Journal of Neuroscience* **17**, 6434–6446.
Lee, Y., Lopez, D., Meloni, E. G., et al. (1994). A primary acoustic startle reflex circuit: role of auditory root neurons and the nucleus reticularis pontis caudalis. *Society for Neuroscience Abstracts* **20**, 1009.
Lee, Y., Lopez, D. E., Meloni, E. G., et al. (1996). A primary acoustic startle circuit: obligatory role of cochlear root neurons and the nucleus reticularis pontis caudalis. *Journal of Neuroscience* **16**, 3775–3789.
Lopez, D. E., Merchan, M. A., Bajo, V. M. and Saldana, E. (1993). The cochlear root neurons in the rat, mouse and gerbil. In: Merchan, M. A. (ed.) *The mammalian cochlear nuclei: organization and function*, pp. 291–301. New York: Plenum Press.
Miserendino, M. J. D. and Davis, M. (1993). NMDA and non-NMDA antagonists infused into the nucleus reticularis pontis caudalis depress the acoustic startle reflex. *Brain Research* **623**, 215–222.
Swerdlow, N. R., Caine, S. B., Braff, D. L., et al. (1992). The neural substrates of sensorimotor gating of the startle reflex: a review of recent findings and their implications. *Journal of Psychopharmacology* **6**, 176–190.
Swerdlow, N. R. and Geyer, M. A. (1998). Using an animal model of deficient sensorimotor gating to study the pathophysiology and treatment of schizophrenia. *Schizophrenia Bulletin* **24**, 285–301.
Walker, D. L. and Davis, M. (1997). Anxiogenic effects of high illumination levels assessed with the acoustic startle paradigm. *Biological Psychiatry* **42**, 461–471.
Walker, D. L. and Davis, M. (1997). Double dissociation between the involvement of the bed nucleus of the stria terminalis and the central nucleus of the amygdala in light-enhanced versus fear-potentiated startle. *Journal of Neuroscience* **17**, 9375–9383.
Yeomans, J. S. and Franklin, P. W. (1996). The acoustic startle reflex: neurons and connections. *Brain Research Reviews* **21**, 301–314.

Steroid Hormone Receptors

M Beato
Centre de Regulació de Genòmica, UPF, Barcelona, Spain
J Klug
University of Giessen, Giessen, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M Beato and J Klug, volume 3, pp 495–504, © 2000, Elsevier Inc.

Historical Development
Steroid Hormone Receptors Form a Subgroup within the Nuclear Receptor Family
Domain Structure of Steroid Hormone Receptors
Individual Family Members
Heterooligomeric Complex and Ligand and DNA Binding
Activation of Transcription
Involvement of the Glucocorticoid Receptor in the Stress Response
Cross-Talk with Other Signal Transduction Pathways

Glossary

Chromatin	A complex of nuclear DNA and basic proteins called histones. Every 146 bp of nuclear DNA is wrapped around a histone octamer into a nucleosome. A 30-bp DNA linker connects one nucleosome to the next, forming a nucleosome necklace.
Steroid hormone receptors (SHRs)	Intracellular transcription factors that are induced by the specific and high-affinity binding of ligands to exert positive or negative effects on the expression of target genes by different mechanisms.

Steroid hormones — Lipophilic hormones derived from cholesterol and synthesized in the adrenal cortex (glucocorticoids, mineralocorticoids, and adrenal androgens), the testes (testicular androgens and estrogen), and the ovary and placenta (estrogens and progestagens).

Transcription factors (TFs) — Nuclear proteins interacting with the regulatory regions of genes, promoters, and enhancers. They can be divided into general transcription factors, forming the transcription initiation complex required for the transcription of all genes, and sequence-specific transcription factors, which bind to specific DNA sequences present in only a subset of genes and carry activation or repression domains that serve to activate or repress transcription.

Steroid hormone receptors (SHRs) are ligand-activated transcription factors and the key mediators of steroid hormone action. Lipophilic steroid hormones pass the cell membrane of target cells by simple diffusion and bind to soluble intracellular SHRs that are complexed to chaperones, such as Hsp90, or to a small fraction of membrane-tethered SHRs. Chaperones are proteins that help other proteins fold and prevent aggregation. The binding of agonistic or antagonistic ligands leads to different allosteric changes of receptors, making them competent to exert positive or negative effects on the expression of target genes by three mechanisms. First, ligand-bound SHR complexes can bind to short, specific DNA sequences in the vicinity of target genes, termed hormone response elements (HREs). The HRE-recruited hormone–receptor complexes are then able to relay activating or repressing signals to the target-gene transcription machinery (**Figure 1a**). Second, the SHR response can be integrated into intracellular signal transduction pathways that transmit an extracellular growth factor signal via mitogen-activated protein kinases (MAPKs) to a nuclear transcription factor that activates a target gene (**Figure 1b**). Third, through protein–protein interactions, SHRs can repress the activity of many genes that are switched on during the stress response (transrepression) (**Figure 1c**).

Historical Development

The gonads and adrenal gland produce five major groups of steroid hormones: estrogens, progestagens, androgens, glucocorticoids, and mineralocorticoids. The effects and clinical symptoms associated with a reduced level of these hormones were known long before the hormone concept was developed. Aristotle, for instance, identified the consequences of castration in humans and birds, and in 1855 Addison described the symptoms of chronic adrenal insufficiency. Moreover, since the nineteenth century it has been known that female sex steroid hormones play a pivotal role in promoting mammary tumor growth, similar to androgens promoting the growth of prostatic neoplasms. Between the tenth and sixteenth centuries, Chinese alchemists developed empirical methods to purify steroids to near homogeneity. In the first half of the twentieth century, it became possible to isolate individual steroids from extracts derived from source tissues by employing biological complementation assays. When tritium and ^{14}C-labeled hormone compounds of high specific activity became available in the late 1950s, it became possible to follow the fate of the steroid hormones up to their target tissues. Thus, it could be shown that the hormone itself, not a metabolite, produced the response via the activation of gene expression mechanisms. The concept that steroid hormones are involved in transcriptional control was also triggered by the observation of ecdysone-induced giant chromosome puffs by Clever and Karlson in 1960. Some 10 years later the groups of Jensen and Gorski had established a two-step model that involved the binding of the hormone to specific high-affinity SHRs within the target cells, followed by the activation of the hormone–receptor complex in order to induce the expression of hormone responsive genes (**Figure 1a**). After another 10 years, the hormone receptors for glucocorticoids and estrogens were cloned and thus became the first molecularly defined transcription factors for RNA polymerase II. At around the same time, the cloning of some SHR targets, such as the human metallothionein gene and the mouse mammary tumor virus genome, led to the identification of the first HREs. Within the last 15 years, many details of the SHR signal transduction pathway, including new mechanisms, have been discovered (**Figures 1b–c**).

Steroid Hormone Receptors Form a Subgroup within the Nuclear Receptor Family

In addition to the hormone receptors for glucocorticoids and estrogens, which were cloned first, receptors for the remaining three major steroid hormone classes have been identified and characterized extensively, as well as a second estrogen receptor and three estrogen-related receptors. All SHRs are characterized by a central DNA-binding domain (DBD) that targets the receptor to the HREs and by a ligand-binding domain (LBD) that is required for switching the receptor's functions. When the v-*erbA* oncogene was cloned, it turned out to also contain a DBD and a LBD that were homologous to the cognate regions of

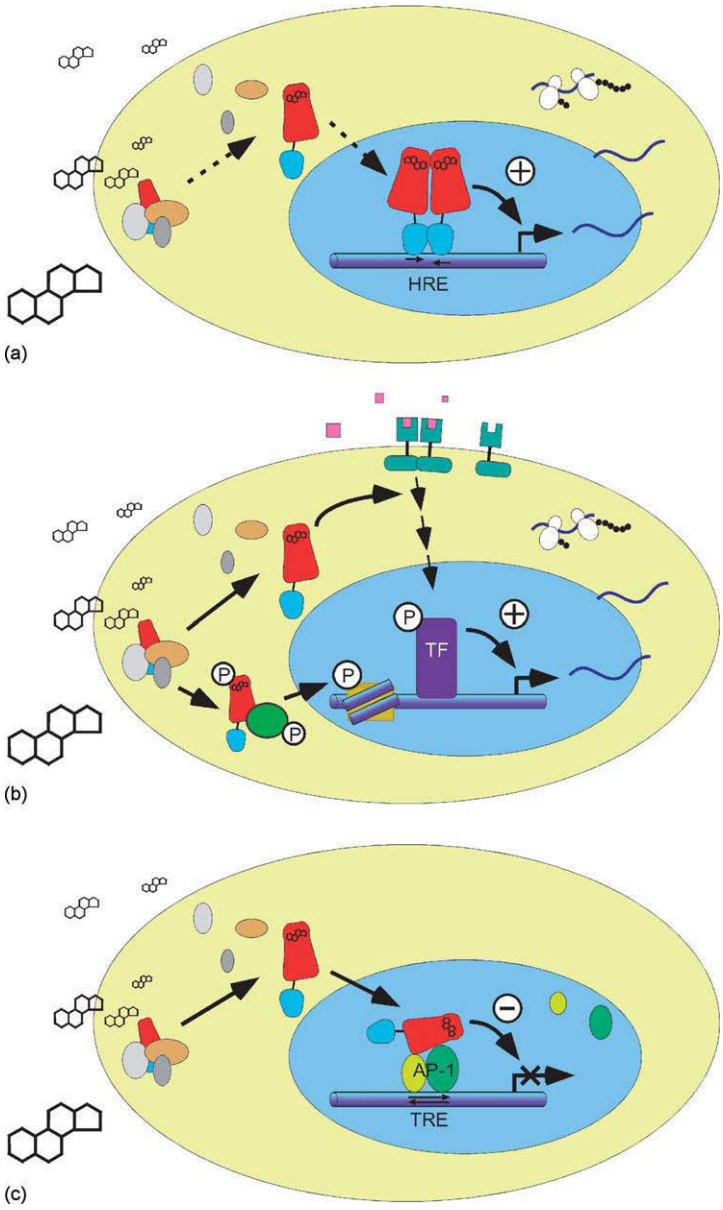


Figure 1 Three implementations of the SHR signal transduction pathway. a, Ligand-bound SHR complexes binding to short, specific DNA sequences in the vicinity of target genes; b, SHR response being integrated into intracellular signal transduction pathways; c, SHRs repressing the activity of genes that are switched on during the stress response. In a, the lipophilic steroid hormones (depicted by the chemical formula showing the central sterane building block of all steroid hormones) pass the cell membrane of target cells by simple diffusion and bind to soluble intracellular SHRs that are complexed to chaperones such as Hsp90 (shown as ellipses with various colors). Binding of the hormone leads to an allosteric change of the receptor, resulting in the dissociation of chaperones (dashed arrow). The liganded SHR-complexes dimerize, bind to hormone response elements (HREs, dashed arrow) in the vicinity of target genes, and transmit activating (or repressing) signals to the target genes' transcription machinery. In b, an extracellular polypeptide growth factor signal (pink squares) is received by the target cell via membrane receptors with kinase activity that dimerize on ligand binding. The dimerized growth factor receptors relay the signal through a cascade of mitogen-activated protein kinases (MAPKs; arrows) to sequence specific transcription factors (TFs) that are induced (or repressed) by phosphorylation (P) to activate the target gene. Hormone-activated SHRs can bypass a growth factor receptor-mediated signal by activating MAPKs and vice versa. In addition, hormone receptors form complexes with cytoplasmic kinases (green ellipsoid) such as ERK and MSK1 that are activated (phosphorylated) on hormone binding and activate (phosphorylate) the receptor. The complex of activated receptor and kinase(s) is recruited to the target promoter where it phosphorylates histones (brown rectangle) as a prerequisite for chromatin remodeling and transcriptional activation. In c, many immunoregulatory genes as well as genes involved in inflammation are positively regulated by the heterodimeric sequence-specific transcription factor activator protein 1 (AP-1) recognizing a TPA response element (TRE) within the promoter of a target gene. Through protein–protein interactions, ligand-activated SHRs are able to repress in trans the AP-1 activity (transrepression).

the known SHRs. Later on, the c-*erbA* locus was identified as the receptor for the thyroid hormone T_3, which is also a hormone found in the nucleus like steroid hormones. The identification of another structurally related receptor for the vitamin A metabolite retinoic acid solidified the concept of a nuclear receptor superfamily. Today, more than 520 different nuclear receptors have been described from more than 120 metazoan species, among them receptors for the known nuclear hormones and a vast number of so-called orphan receptors with no, or unknown and possibly novel, ligands. In order to avoid the coining of several trivial names for the same gene, an official nomenclature system has been in use since 1999 (NUREBASE). The SHRs and steroid hormone-like receptors form the nuclear receptor subfamily 3 (NR3), consisting of three groups: A, B, and C (Figure 2a). NR3 includes two receptors for estrogens (ERα and ERβ, or NR3A1 and NR3A2), three estrogen-related orphan receptors (ERRα, ERRβ, and ERRγ, or NR3B1, NR3B2, and NR3B3), and one receptor each for the other major steroid hormone classes: glucocorticoids (GR, or NR3C1), mineralocorticoids (MR, or NR3C2), progestagens (PR, or NR3C3), and androgens (AR, or NR3C4).

Domain Structure of Steroid Hormone Receptors

All SHRs are modular proteins composed of distinct regions, as shown in Figure 2b (top line). When the chicken estrogen receptor cDNA became available, comparison with the human estrogen and glucocorticoid receptor sequences led to the region A–F nomenclature, which is used for all members of the whole nuclear receptor superfamily today. From numerous functional and structural analyses, it became clear that these distinct regions correspond to functional and structural units called domains. Region C (the DBD) and region E (the LBD) display a high degree of sequence conservation, whereas no significant conservation is detected between paralogous SHRs for the regions A/B, D, and F.

Region D is considered a flexible hinge region between the DBD and LBD. Its very N-terminus is an integral part of the DBD and is involved in DBD dimerization. The F region is essential for hormone binding in the PR, GR, and AR but not in the ERα. This domain is also important for the discrimination between agonistic and antagonistic hormone ligands. The A region is highly conserved between chicken and human estrogen receptors, but this distinction is much less clear in the other steroid receptors. Therefore, regions A and B are combined into an A/B region in most cases.

Regions C and E are not only responsible for DNA and ligand binding, respectively, but they do encode other functions as well (Figure 2b, bottom). The intracellular distribution of steroid receptors is the result of nuclear-cytoplasmic diffusion and cytoplasmic-nuclear shuttling. At equilibrium, the majority of steroid receptors are in the nucleus due to the presence of nuclear localization signals (NLS), which are believed to be required for nuclear pore recognition. A constitutive NLS is located at the border of regions C and D, whereas a second NLS in the LBD is ligand-dependent. Intracellular localization is different for GR and MR because hormone-induced nuclear translocation has been reported in these cases.

Most SHRs bind as homodimers to their cognate response elements within the promoters of target genes, which is a specific characteristic of this subfamily within the nuclear receptor superfamily. It is also well known that the predominant form of the GR is monomeric in solution and that dimerization occurs only after binding to an HRE, accounting for the cooperativity observed during DNA binding. Responsible for dimerization are a weak dimerization domain encoded within the DBD and a strong dimerization interface provided by the LBD.

Ligand binding confers transcriptional competence on to SHRs that is exerted in most receptors by two independent transactivation functions: a constitutively active one in the A/B region located close to the DBD, referred to as activation function 1 (AF-1; also called τ-1 or enh-1), and a ligand-inducible activation function in the LBD, called AF-2. AFs connect the receptor to the transcription apparatus via direct interactions with basal transcription factors, sequence-specific transcription factors, and/or coregulators.

Individual Family Members

The human estrogen receptor (now called hERα, Figure 2a) was cloned from a cDNA expression library produced from the breast cancer cell line MCF-7. A second estrogen receptor, ERβ, has been characterized that is very similar to ERα in terms of structure and function but also shows subtle and important functional differences. Both receptors bind the ligands estradiol, diethylstilbestrol, estriol, and estrone with high affinity. cDNA clones for hERRα and hERRβ were isolated from a human testis cDNA library using the hERα DBD as a probe. Mouse ERRγ was identified by yeast two-hybrid screening, using the transcriptional coactivator GR interacting protein 1 as bait. It is still unclear whether the ERRs are true orphan receptors in search of a ligand or are ligand-independent.

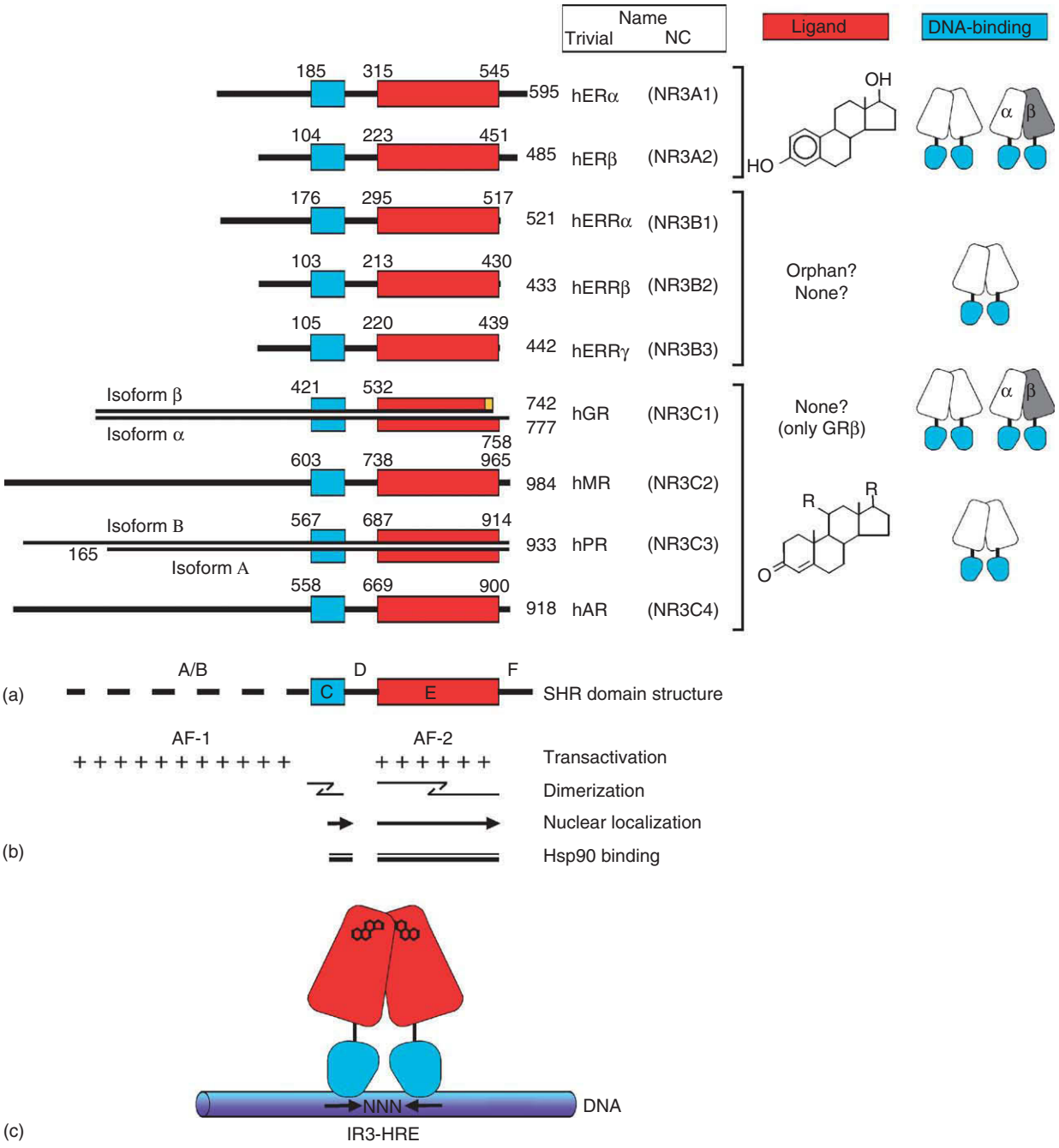


Figure 2 The human steroid hormone receptor family. a, Structure of all family members, cognate ligand, and DNA-binding mode; b, SHR domain structure and structure–function relationships; c, an SHR homodimer bound to its cognate hormone response element (HRE). In a, the numbers refer to amino acid positions. Highlighted are the DNA-binding domain (DBD; shown in blue) and the ligand-binding domain (LBD; shown in red). The A/B and F domains are drawn to scale, and the name of each receptor is indicated (trivial and name according to nomenclature scheme). In c, the HRE is of an inverted repeat (IR) type (palindrome), and half-sites (arrows) are separated by three unspecified (N) nucleotides (IR3-HRE); the hormone ligand is indicated by a sterane formula. AF-1 and-2, transcription activation functions 1 and 2; AR, androgen receptor; ER, estrogen receptor; ERR, estrogen-related receptor; GR, glucocorticoid receptor; h, human; MR, mineralocorticoid receptor; PR, progesterone receptor.

For the GR, two different classes of cDNAs have been described in humans, termed hGRα and hGRβ, which are the result of alternative splicing from a single gene transcript. Both isoforms are identical up to amino acid 727 and then diverge, with hGRα being slighly larger (777 amino acids) than hGRβ (742 amino acids). In contrast to hGRα, which is commonly referred to as the bona fide hGR, hGRβ was long dismissed as a cloning artifact but has now been shown to be expressed at modest but varying levels

in a range of tissues. Each variant mRNA is translated from at least eight initiation sites into multiple GRα and possibly GRβ isoforms, amounting to a multitude of GR monomers and probably a plethora of homo- or heterodimers of unknown function. hGRβ does not bind hormones, is transcriptionally inactive, and therefore acts as a ligand-independent negative regulator of glucocorticoid action in transfection experiments. However, the β isoform is not conserved across species, and functional evidence that implicates hGRβ in glucocortioid insensitivity is still missing. hGRα shows a high affinity for the artificial glucocorticoid dexamethasone, moderate affinity for the physiological steroids cortisol and corticosterone, and low affinity for mineralocorticoids and progesterone. In contrast to all other SHRs, GRα dimerizes only weekly.

The progesterone receptor was the first SHR shown to exist in two common isoforms generated by differential promoter use. One promoter initiates transcription at positions +1 and +15 of the gene, which gives rise to the longer isoform B (**Figure 2a**). The second promoter initiates hPR transcripts between +737 and +842, encoding the 164-amino-acid residues shorter hPR form A. Only PR-B harbors a third activation function (AF-3) at its specific N-terminus, which functions promoter- and cell-specificly. Both isoforms, therefore, display differential target gene specificity. First regarded as an exception in the family, differential promoter use and alternative splicing are now considered the rule for all members. Both PR isoforms show a high affinity for the natural ligand progesterone and the synthetic agonist R5020. Alternative splicing can also generate truncated forms of PR that are found in breast cancer.

The human mineralocorticoid receptor (hMR) and androgen receptor (hAR) were cloned 2–3 years after the GR was described. The hAR exists in two isoforms, hAR-A and hAR-B, which are structurally analogous to the two hPR isoforms. In contrast to hPR-A/B, hAR-A is expressed at substantially lower levels than the B form and its contribution to androgen action is not known. The hAR binds the two naturally occurring ligands dihydrotestosterone and testosterone with high affinity, whereas the hMR shows high and equivalent affinity for aldosterone, physiological corticosteroids, and progesterone.

Heterooligomeric Complex and Ligand and DNA Binding

Heterooligomeric Complex

In 1965, the groups of Jensen and Gorski showed that specific and high-affinity binding sites for estrogen existed in large heterooligomeric complexes migrating with 9S on low-salt sucrose gradients. This non-DNA-binding complex can be transformed into a more compact and protease-resistant 4S DNA-binding form by high salt and temperature or hormone. The 9S heterooligomeric receptor complexes were later shown to be composed of the two chaperones – heat shock protein (Hsp)90, recently identified as a member of the GHKL family of ATPases, and Hsp70 – plus a number of co-chaperones.

Hsps form an evolutionarily very conserved group of stress-inducible proteins. Contrary to their name, Hsps are strongly induced not only by heat shock, but are also expressed constitutively at physiological temperatures at significant levels. Hsps belong to a larger family of chaperones that are known to stabilize the native structure of other polypeptides by controlled binding to them.

Most investigations have been performed with the heterooligomeric GR complex that contains a dimer of Hsp90 and a substoichiometric amount of Hsp70 and one or two molecules of p23. The last two are required for the assembly of the final heterooligomeric complex. The crystal structure of full-length yeast Hsp90 in complex with an ATP analog and the yeast p23 homolog has been solved recently. For PR and AR, the LBD is sufficient for an interaction with Hsp90 that also requires ATP. In case of the GR, a seven-amino-acid segment overlapping the N-terminus of the LBD, which is located on the surface of the LBD at the opening of the ligand-binding pocket, is sufficient and required for Hsp90 binding. In case of the ER, in addition to the LBD, part of the hinge region D stabilizes the Hsp90–ER complex. From the known crystal structures of a number of LBDs of nuclear receptors and several biochemical studies, it becomes clear that Hsp90 recognizes two hydrophobic surfaces at opposite sides of the LBD located in the vicinity of the ligand-binding pocket. The large distance between the two hydrophobic surfaces probably accounts for the fact that two Hsp90 molecules are bound to SHRs. The LBD of the receptor must be in a complex with Hsp90 in order to have high ligand-binding activity. Hsp90 binding seems to be a characteristic of SHRs because other nuclear receptors, such as the retinoid, thyroid hormone, and vitamin D_3 receptors, do not associate efficiently with these Hsps.

The complexes also contain one of a number of tetratricopeptide (TPR) domain immunophilins that bind to a TPR acceptor site at the C-terminus of Hsp90. Immunophilins are characterized by their peptidylprolyl isomerase (PPIase or rotamase) domains and can function as chaperones like Hsps. SHR complexes contain one of three immunophilins (FKBP51,

FKBP52 or CyP-40) or PP5, a protein phosphatase that contains both a TPR domain and a PPIase homology domain. The PPIase domain of immunophilins is a binding site for immunosuppressant drugs such as FK506 or cyclosporine A that inhibit PPIase activity on binding.

In hormone-free cells, steroid receptors shuttle between the cytoplasm and nucleus. The distribution may be predominantly cytoplasmic (GR) or nuclear (ERα). Steroid-dependent transformation of the GR triggers its rapid translocation to the nucleus, possibly involving movement along microtubules driven by the motor protein dynein. In wheat germ extracts, it was shown *in vitro* that wheat immunophilins can link the GR–Hsp90 heterocomplex to cytoplasmic dynein via conserved TPR and PPIase domain interactions. Immunoadsorption of GR from cell lysates of L cells coimmunoadsorbs Hsp90, the dynein intermediate chain, FKBP52, PP5, and dynamitin (a subunit of the dynein-associated dynactin complex). If the microtubules are stabilized by taxol, α-tubulin is coimmunoadsorbed as well.

All these data are derived from *in vitro* experiments with cell extracts in which the cellular microenvironments are destroyed, so that the observed complexes may be artificial. But a number of *in vivo* observations indicate that the heterooligomeric receptor complex is also of functional importance. Increased levels of FKBP51 have been reported to be the common cause of glucocorticoid resistance in three New World primates, due to delayed translocation of the GR to the nucleus. Likewise the overexpression of FKBP51 in COS cells reduces the hormone binding activity of GR and hence its transcriptional activity. In yeast cells, the overexpression of FKBP51 has no effect, whereas FKBP52 enhances GR-dependent transcription that is dependent on interaction with Hsp90 and PPIase activity. It could be shown that the binding of FKBP52 to Hsp90 potentiates glucocorticoid-dependent reporter gene activity in yeast and mammalian cells that is dependent on PPIase activity and increases the hormone binding affinity of the receptor. Moreover, differential binding of FKBP51 and-52 to dynein corresponds to a differential effect on nuclear translocation. The co-expression of dynamitin and the PPIase domain, as well as PPIase domain-swapping experiments, demonstrated the relevance of the dynactin complex for receptor function. Hsp90 and the co-chaperone p23 have been shown to colocalize *in vivo* at genomic HREs in a hormone-dependent manner. p23 seems to be involved in disrupting receptor-mediated transcriptional activation *in vivo* and *in vitro* and, thus, it could be involved in terminating the hormone response.

Therefore receptor-associated chaperones and co-chaperones can no longer be viewed as simple repressors that have to be removed by hormone in order to transform the non-DNA-binding heterooligomeric complex into a DNA-binding entity; they are, instead, viewed as chaperone complexes that accompany all steps in the life of these client proteins: folding, assembly of quaternary structures, intracellular transport, and targeting for proteolytic degradation.

Ligand Binding

Although not proven formally, it is believed that the lipophilic steroid hormones, as well as synthetic compounds with agonistic or antagonistic effects, enter the target cell by simple diffusion and bind to SHRs within the heterooligomeric complex. For hERα and hPR, the three-dimensional structure of their ligand-binding domains complexed with 17β-estradiol and the antihormone raloxifene or progesterone, respectively, have been determined. Like other known LBDs of nuclear receptors, the LBDs of SHRs are folded into a three-layer antiparallel α-helical sandwich that creates a wedge-shaped molecular scaffold with the ligand-binding cavity at the narrower end of the domain (**Figure 3**). This cavity is completely partitioned from the external environment and is closed by helix 12 of the LBD, operating as a lid after the hormone has entered the binding pocket. The relocation of this amphipathic helix 12 over the hormone-binding site generates a new surface that allows coactivators to bind to the LBD, thereby mediating the activity of AF-2, which forms the core of helix 12. The LBD dimerization interfaces are formed by α-helices that line up and/or intertwine and differ slightly between ERα and PR.

DNA Binding

The receptors for progestagens, glucocorticoids, mineralocorticoids, and androgens bind to the same HRE that is composed of two half-sites (AGAACA), which are arranged as inverted (palindromic) repeats (IR) with the half-sites being separated by three nucleotides (IR-3). The HRE half-site recognized by both ERs (AGGTCA) deviates in two nucleotides from this DNA sequence (**Figure 3c**). One HRE half-site is recognized by one receptor monomer, and the homodimer binds in a head-to-head orientation to an IR-3 element. The non-ERs recognize their half-site mainly by a hydrophobic interaction with the methyl group of the thymine in position 4 of the complementary strand (AGA*A*CA) that is not present in the DNA sequence recognized by the ER. The half-site spacing of 3 bp is such that an SHR homodimer binds both half-sites on the same face of the DNA. In addition to IR-3 elements, only the AR is able to selectively

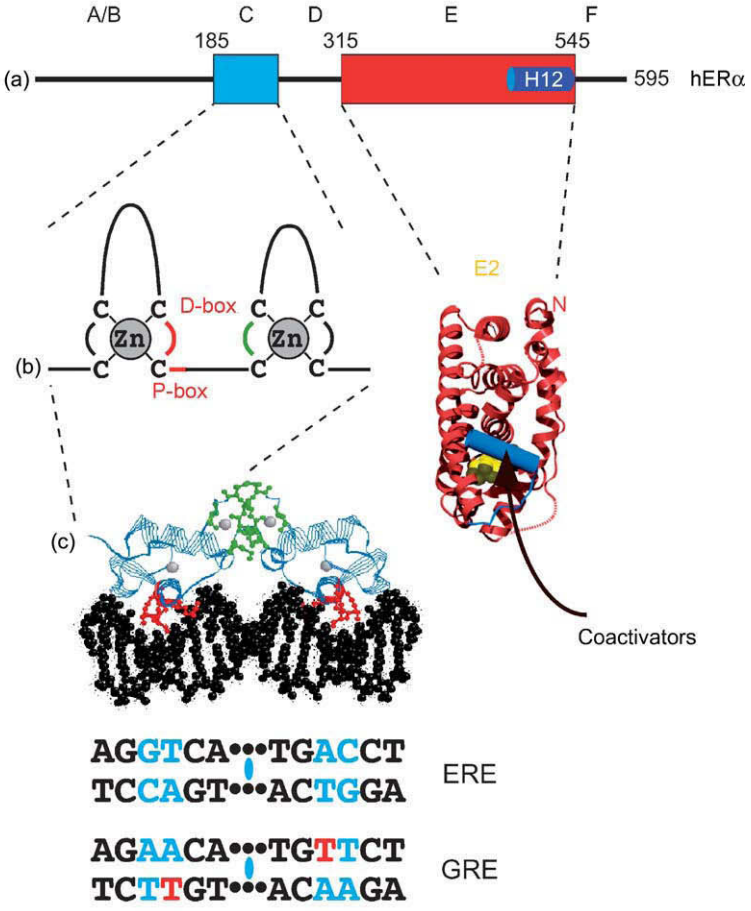


Figure 3 Structure of the human estrogen receptor α. a, Functional domains (A/B to F) of hERα at the primary structure level, with the DNA-binding domain (DBD) and the ligand-binding domain (LBD) highlighted; b, an exploded view of the secondary structure of the hERα DBD; c, crystal structure of the hERα DBD bound to DNA, followed by the sequences of the consensus estrogen response element (ERE) and a consensus glucocorticoid response element (GRE); d, crystal structure of the hERα LBD complexed with 17β-estradiol (E2). In a, The numbers refer to amino acid positions. DBD is highlighted in blue and the LBD in red. The A/B and F domains are drawn to scale. In b, the secondary structure of the DBD is characterized by two steroid hormone receptor-specific zinc fingers. Four cysteines each tetrahedrally coordinate one zinc ion (grey). The Proximal box (P-box) responsible for specific DNA recognition is shown in red; the Distal box (D-box) mediating DBD dimerization is shown in green. In the crystal structure in c, the four amino acid side chains of the P-box (red) are shown interacting with DNA bases (black). These residues are part of an α-helix responsible for specific DNA-sequence recognition that is positioned within the DNA's major groove (perpendicular to the paper plane). A second amphipathic α-helix can be seen to cross the recognition helix with the red P-box residues at a right angle. Between the two helices, the green D-box residues promoting DBD dimerization are located. Underneath the DNA, the sequences of the consensus ERE and a consensus GRE are shown. Both half-sites are organized as an inverted repeat with 3-bp spacing (IR-3). The symmetry axis is indicated by a blue ellipse. The bases that differ between ERE and GRE are highlighted in blue, and the thymine important for the discrimination between both response elements is highlighted in red. In the crystal structure in d, E2, a natural ligand, is shown in yellow. Decisive for a productive interaction with coactivators is the proper positioning of helix 12 (H 12) over the ligand-binding pocket. (We thank R. Hubbard and A. Pike, University of York, UK, for providing this part of the figure.)

recognize direct repeats of the two half-sites separated by three nucleotides (ADR-3 elements). In contrast to other nuclear receptors binding to DR elements, the AR homodimer also binds in a head-to-head orientation to an ADR-3 element.

The DBD of SHRs comprises approximately 80 amino acids encoded by region C plus some 14 N-terminal amino acids of region D. Domain C contains two patterns that are reminiscent of, but clearly distinguishable from, the zinc-finger motifs first observed in the *Xenopus laevis* transcription factor IIIA. The SHR zinc fingers are also able to tetrahedrally coordinate a zinc atom, but are of the type (Cys2-Cys2). Only very few amino acids, termed the proximal (P-)box, within the first SHR zinc finger are responsible for specific recognition of the cognate HRE. Another set of amino acids, called the distal (D-)box within the second SHR zinc finger, forms the weak dimerization interface of the DNA-binding domain (Figure 3b).

Activation of Transcription

To regulate transcription, agonist-liganded SHRs must talk to the general transcription machinery. This can be achieved either by a direct contact between SHRs and the general transcription factors (GTFs), which are building up the transcription preinitiation complex at the transcription start point, or by means of other cellular factors that do not bind to DNA directly but are recruited to the promoters by the SHRs via protein–protein interactions. Depending on the effect of these coregulators on gene expression, they can be divided into coactivators, which promote transcription mediated by SHRs, and corepressors, which suppress SHR-dependent gene transcription. Coregulators, other sequence-specific transcription factors, and chromatin factors all function as interpreters between SHRs and the general transcription machinery.

Coactivators

Coactivators were originally called transcriptional intermediary factors because they are supposed to bridge between DNA-bound sequence-specific transcription factors and GTFs. The suspicion that SHRs may require coactivators for activation was derived from the squelching phenomenon – the observation that an excess receptor can inhibit its own transactivation as well as transcription by other transactivators. One possible explanation for this behavior is that additional factors, which are present in limited amounts and required for transactivation, are trapped by excess SHRs in unproductive complexes. Conversely, by overexpressing coactivators, it is possible to reactivate a target promoter that has been squelched by excess SHR.

Today, more than 20 of these coactivators are known that were quickly identified within a decade. An important and well-characterized class of coactivators is the steroid receptor coactivator (SRC, or p160) family of coactivators, consisting of its three members SRC-1, SRC-2 (also known as transcription intermediary factor 2 or glucocorticoid receptor-interacting protein 1), and SRC-3. The three related proteins have partial overlapping specificity for hormone receptors and cell type. SRCs enhance hormone-dependent transcriptional activities of SHRs without altering the basal activity of a target promoter. The biological functions of SRCs have been revealed by gene-disruption experiments in mice. Whereas SRC-1-deficient mice display normal growth and fertility but partial resistance to a number of steroid hormones, SRC-2 knockout mice have reduced fertility due to gonadal failure and are leaner because the ratio of white and brown adipose tissue is decreased. The inactivation of SRC-3 has more dramatic effects and results in growth retardation and smaller body size as well as resistance to growth hormones (GHs) and estrogen. Like SHRs, coactivators are modular proteins with common functional domains such as one that possesses acetyltransferase activity. For the binding to agonist-liganded SHRs, one or several copies of an α-helical LXXLL signature motif, also called a nuclear receptor (NR) box, are required. After the binding of an agonistic ligand to the SHR, helix 12, containing the AF-2 core of the receptor, is relocated, exposing a hydrophobic cleft and an adjacent ionic clamp as the principal determinants for the interaction with an NR box.

In vivo coactivators are organized into complexes that are poised for interaction with SHRs in response to appropriate stimuli. Many of the characterized complexes share subunits, suggesting that a directed or cyclical assembly of coactivator complexes takes place. Current data support a model in which the initial binding of ligand results in the dissociation of corepressors and recruitment of chromatin remodeling machines such as SWI/SNF. The binding of SRCs and other coactivators such as CBP/p300 then results in local acetyltransferase activity and the opening of chromatin in the vicinity of the target promoter. C-terminal SRC-domains can recruit other coactivators, such as the coactivator-associated arginine methyltransferase-1 (CARM1) and cAMP responsive element binding protein (CREB)-binding protein (CBP), and the related protein p300 for complementing and potentiating histone acetyltransferase activity, respectively.

Coactivators are subject to a number of posttranslational modifications for modulating coactivator function. Acetylation and ubiquitination appear to alter the half-life of molecules, whereas phosphorylation is essential for the optimal activity of coactivators within the cell. Phosporylation of SRCs occurs in response to a plethora of stimuli, including epidermal growth factor, steroid hormones, cytokines, and increased intracellular cAMP. In turn, multiple kinases, such as MAPKs and IκB kinases, that are involved in a number of cellular signaling pathways are responsible for a distinct phosphorylation pattern on SRC-3, regulating its activity and, more important, its specificity. In contrast to SHRs, SRCs have a broad tissue distribution and coactivate an ever-expanding number of transcription factors. Therefore, coactivators, like SRCs, might serve as integrators of multiple afferent stimuli in order to produce an appropriate cellular response.

Cyclin D1, the regulatory subunit of a holoenzyme that phosphorylates and inactivates the retinoblastoma protein tumor suppressor and promotes progression through the G1-S phase of the cell cycle, has

been identified as a coactivator for ERα recently. The *cyclin D1* gene is frequently overexpressed in human breast cancer and is capable of inducing mammary tumorigenesis when overexpressed in transgenic mice. Cyclin D1 binds to the N-terminal half of the ligand binding domain of ERα and enhances ligand-induced recruitment of ERα to an estrogen responsive element (ERE). More important, cyclin D1 antagonizes the repression of ERα activity mediated by the tumor suppressor gene product BRCA1, which is functioning as a corepressor of ERα in this context. Thus cyclin D1 and BRCA1 constitute an antagonizing pair of coregulators for ERα.

Corepressors

For some members of the nuclear receptor family, such as the thyroid hormone receptor (TR) and the retinoid acid receptors (RARs), that form heterodimers with the retinoid X receptor (RXR), it is well known that they repress basal transcription in the unliganded state. It has been shown that this silencing effect is mediated by corepressors that were called nuclear receptor corepressor (N-CoR) and silencing mediator for retinoid and thyroid hormone receptors (SMRT). N-CoR and SMRT associate with DNA-bound TR–RXR or RAR–RXR heterodimers in the absence of hormone but not in its presence. SMRT and N-CoR are multiprotein complexes that exhibit histone deacetylase activity. Both corepressors are also involved in steroid hormone action because it was demonstrated that the overexpression of N-CoR and SMRT represses the partial agonistic activity of both ERα bound to the antihormone tamoxifen and of PR bound to the antagonist RU486. Recognition by transcriptionally incompetent SHRs of corepressors is mediated by short amphipathic helical signature motifs, called CoRNR boxes, which are similar to the NR boxes of coactivators. Corepressor activity is mediated by histone deacetylation. Because corepressors lack intrinsic deacetylation domains, they recruit histone deacetylases for this task. Apparently, corepressors and coactivators are functionally similar proteins, so it is not surprising that a corepressor can also turn into a coactivator, as shown for N-CoR and RAR.

Recently, the tumor suppressor gene product BRCA1 has been shown to be a corepressor of ERα. BRCA1 is implicated in DNA repair mechanisms, and germ-line mutations in the BRCA1 gene predispose women to ovarian and breast cancer. BRCA1 inhibited estradiol-induced endogenous gene expression at levels of BRCA1 that are within the physiological range. Moreover, BRCA1 was shown to bind to ERα directly *in vitro* and the binding of BRCA1 to an estrogen responsive gene was reduced by estradiol *in vivo*. BRCA1 appears not to act on specific estrogen responsive genes or gene subsets but blocks the ability of estradiol to alter the expression of the vast majority of all estradiol responsive genes in MCF7 cells. BRCA1 downregulates the coactivator p300, and excess p300 rescues BRCA1 repression of ERα like cyclin D1 can. The physical association of cyclin D1 with ERα may exclude the direct association of BRCA1 with the receptor or alter the conformation of BRCA1 and/or the binding of other BRCA1-associated proteins and thereby contributes to activation at an ERE.

Altogether, SHRs can exist in a multitude of conformations, depending on the nature of bound ligand. The various interaction surfaces are mainly determined by helix 12 and most likely by the C-terminus beyond helix 12, the F region. The transactivation by SHRs appears to be not only a simple net activation but the sum of relief from repression by corepressors and activation by coactivators. The switch from the repressed to the activated state is inititated by the hormone ligand through an allosteric change in SHR structure followed by an intricate interplay of coregulators.

Interactions with Sequence-Specific Transcription Factors

SHRs control the activity of natural promoters through positive and negative interactions with sequence-specific transcription factors. Particularly interesting is an interaction that represses the activity of both partners, as exemplified in the case of hGR and the heterodimeric transcription factor AP-1. Similar interactions have been described between GR and the p65 subunit of the transcription factor nuclear factor (NF-)κB and the transcription factor GATA-1. From experiments with numerous SHR mutants, it can be concluded that repression is most likely mediated by a protein–protein interaction with an SHR monomer that requires the N-terminal portion of the receptor. Because most immunoregulatory genes as well as genes involved in inflammation are positively regulated by the transcription factors AP-1 and NF-κB, it is conceivable that the immunosuppressive and anti-inflammatory activities of glucocorticoids are mediated through the inhibition of AP-1- and NF-κB-mediated transactivation by GR (**Figure 1c**).

A novel biological action of GR was revealed by constructing mice with the hepatocyte-specific loss of GR that have a severe reduction in body weight due to decreased bone length and size of internal organs, very reminiscent of mice lacking the signal transducer and activator of transcription (STAT)5. In both mice strains, growth hormone (GH) signaling via

the JAK–STAT pathway is impaired. GR directly interacts with both phosphorylated and unphosphorylated forms of STAT5, and both proteins are detectable at promoters of GH-regulated genes *in vivo,* which is essential for GH signaling. As shown by dimerization defective GR mutant (GR^{dim}) mice, GH signaling is independent of GR DNA binding and most likely mediated by protein–protein interaction with STAT5 transcription factors.

Interaction with Chromatin Factors

The interaction of SHRs with DNA, the general transcription machinery, coregulators, and sequence-specific transcription factors takes place in the nucleus with its DNA compacted into chromatin. Genetic analyses have demonstrated a widespread involvement of chromatin structure in gene regulation in general. For the GR, it was shown that components of the so-called SWI–SNF complex, a set of pleiotropic transactivators that counteract the repressive functions of chromatin, are required for transactivation in yeast. In human cells lacking brahma (hBrm), the homolog of the SWI–SNF complex subunit SWI2, transactivation by GR is weak and can be enhanced selectively by the ectopic expression of hBrm. Like SWI2 in yeast, hBrm is part of a large multiprotein complex that mediates the ATP-dependent disruption of a nucleosome. A direct interaction between nuclear receptors and components of hBrg1, a brahma-related ATPase complex, has been demonstrated, and chromatin immunoprecipitation experiments indicate that the hormone receptors recruit these ATP-dependent chromatin remodeling complexes to their target genes. In the case of the MMTV promoter, activation by glucocorticoids and progestins depends on the correct organization of the promoter sequences in nucleosomes so that some HREs of a larger cluster get exposed. The binding of the receptor to the exposed HREs leads to the recruitment of ATP-dependent chromatin remodeling activities that displace H2A–H2B dimers from the promoter and allow the binding of the ubiquitous transcription factor NF1 and the recruitment of more receptor molecules, followed by coactivators and the transcription preinitiation complex. Optimal activation requires the inclusion of linker histones that help to position nucleosomes and generate a more suitable platform for receptor binding.

Apart from the SWI/SNF complex and related ATP-dependent chromatin remodeling machines, increasing evidence has implicated the modification of histones in the regulation of gene transcription. Histones are composed of a histone fold domain, involved in wrapping the DNA in nucleosomes, and an N-terminal tail rich in lysine that protrudes out of the nucleosome. The acetylation of tail lysines reduces the affinity of histone tails for DNA greatly and is believed to render DNA more accessible to transcription factors while still maintaining a nucleosomal architecture. Moreover, some of the identified coactivators for SHRs, such as CBP/p300 and SRC1, as well as the largest subunit of the pivotal GTF TFIID, $TAF_{II}230/250$, turned out to be histone acetyl transferases (HATs). All of these observations led to the current two-step model for transcriptional activation by SHRs that is built on the hormone-mediated recruitment of coactivators and other transcription factors with HAT activity, resulting in the local destabilization of repressive histone–DNA interactions, followed by direct or indirect interactions with the basal transcription machinery. Histone arginine methyl transferases, such as CARM1, are also implicated in gene activation by steroid hormones. and recent evidence suggests a participation of kinases that phosphorylate core and linker histones. Altogether, these new findings emphasize the importance of chromatin remodeling in gene regulation by steroid hormones.

Recently, a detailed kinetic analysis by chromatin immunoprecipitation (ChiP) analyses of the assembly of transcription initiation complexes on the estrogen responsive promoter of the pS2 gene has revealed a sequence of cycles of formation of multiprotein complexes involving protein-modifying enzymes and ATP-dependent chromatin remodeling complexes, followed by ubiquitin/proteasome-mediated degradation of the receptor with a cycling amplitude on a time scale of minutes. The dynamic nature of transcriptional activation has also been assessed by single live-cell imaging of transcription factors tagged with fluorescent proteins. Contrary to the ChiP results, fluorescence recovery after photobleaching analyses have shown that SHRs and interacting coregulators rapidly exchange on large artificial HRE arrays on a time scale of seconds. In attempting to reconcile these contradictory results, we have to consider that ChIP is a composite freeze-frame of promoter occupancy in millions of cells, whereas live-imaging focuses on the chronology of events in a single cell.

Involvement of the Glucocorticoid Receptor in the Stress Response

Studies of human patients and adrenalectomized animals have shown that glucocorticoids have a plethora of functions, the most important being the stimulation of gluconeogenesis; decrease of glucose uptake by peripheral tissues; increase of glycogen deposition; increase of lipolysis; stimulation of protein and nucleic acid synthesis, predominantly in the liver; suppression of immune and inflammatory responses;

inhibition of fibroblast proliferation and extracellular matrix deposition; maintenance of normal cardiovascular function and blood pressure; and regulation of calcium absorption. Many of these functions should prevent an overreaction of defense mechanisms to stress and be conducive to coping with the stress response. In line with this hypothesis, the synthesis and release of glucocorticoids are elevated during stress, and people with chronic adrenal insufficiency show "general languor and debility, remarkable feebleness of the heart's action, irritability of the stomach and a peculiar change in the color of the skin," as Addison described it in 1855. Although none of these symptoms is life-threatening, patients with Addison disease cannot withstand stress and are prone to suffering an acute adrenal crisis that can lead to shock or coma.

Until recently, it was believed that all these glucocorticoid effects were of a genomic nature, mediated by its receptor binding to glucocorticoid response elements (GREs) in the target genes. Given the multiple important functions of glucocorticoids, it was therefore not surprising to learn that GR-deficient mice, which were generated using gene knockout techniques, are not viable and die shortly after birth.

However, the cause of death – respiratory failure due to lack of lung maturation – does not seem to depend on genomic effects mediated by GREs. Contrary to GR-deficient mice, homozygous mice that express a mutated GR that is not able to dimerize appear almost normal and healthy under standard laboratory conditions. This dimerization-defective GR (GR^{dim}) can no longer bind efficiently to palindromic GREs, and these mutant mice do not respond to the administration of glucocorticoids by the induction of well-characterized hormone responsive genes. Clearly, GRE-mediated gene activation is not required for development and survival. On the other hand, GR^{dim} mice show a selective impairment of spatial memory in the water maze. It is known that the stress that is associated with a learning task – here orientation in a maze – leads to the activation of GRs in the brain, which is required for facilitated storage of the acquired information. Therefore, the facilitating effects of corticosterone on spatial memory seem to depend on DNA binding of the GR.

There is also increasing evidence that altered GR signaling is an important cause of the pathogenesis of major depression. The most consistent biological finding in patients with depression is hyperactivity of the hypothalamic-pituitary-adrenal (HPA) axis. Mutant mice with a 50% GR gene dose reduction have a disinhibited HPA axis like human patients and show increased helplessness after stress exposure, a behavioral correlate of depression in mice. In contrast, mice that overexpress GR demonstrate a reduced helplessness response after stress challenge. Also, human carriers of two polymorphisms in the GR gene had an increased risk of developing a major depressive mode.

Cross-Talk with Other Signal Transduction Pathways

Steroid hormones can either act via their cognate receptors that alter gene expression after binding to HREs in target genes (termed genomic or nuclear action) or act via membrane-bound or -localized SHRs or non-SHRs (termed nongenomic or cytoplasmic action). Whereas genomic action is slow and takes 30–60 min, nongenomic action is rapid and takes a couple of seconds or minutes. We now briefly summarize the important findings on nongenomic actions that involve the classical steroid receptors.

The existence of such nongenomic actions became apparent in experiments with GR knockout mice that die postnatally and with mice that express a dimerization-defective mutant GR and are viable. The dimerization-defective GR shows impaired GRE-dependent transactivation while functions that require cross-talk with other transcription factors remain intact. Moreover, many studies employing different techniques and assays have shown that a small pool of estrogen (some 2%) and ARs are associated with the plasma membrane. More specifically, both receptors have been identified in the caveolae, together with a number of signaling molecules. One of the actions of steroid receptors thought to take place in the caveolae is the estrogen-mediated rapid increase in endothelial nitric oxide synthase (eNOS) activity. The 110-kDa caveolin-binding protein striatin has been identified as a molecular interaction hub that organizes an ERα–eNOS membrane signaling complex engaging also the G-protein $G_{\alpha i}$. The disruption of the ERα–striatin interaction abolishes the rapid activation of MAPK, Akt kinase (also called protein kinase B), and eNOS by estrogen, whereas genomic signaling is not affected. But other proteins such as caveolin-1, modulator of nongenomic action of estrogen receptor (MNAR), Shc, or insulin-like growth factor-1 receptor are also implicated in tethering ERα to the inner face of the plasma membrane bilayer. Moreover, the steroid-independent palmitoylation of a cystine residue within the E domain has been observed. Palmitoylation facilitates caveolin-1 binding as well as membrane localization. Interestingly, A/B and C domain-deleted ERα still localizes to the membrane and signals to extracellular receptor kinase (ERK) like the wild-type receptor. Current evidence suggests that ERα membrane localization is required for kinase activation.

A well-known rapid response of steroid receptors is the activation of the c-Src–p21ras–ERK2 pathway by ligand-bound AR, ERα (but not ERβ), GR, MR, and PR-B. For all five SHRs MAPK activation is abrogated in the presence of antihormones, but this is less clear for MR. Although various mechanisms have been proposed, the interaction of SHRs with the kinase c-Src is an important aspect of cross-talk. Whereas AR interacts with the SH3 domain of c-Src, ERα interacts with the SH2 domain. AR activation of the ERK pathway is more effective if, in addition to AR binding to the SH2 domain, ERα simultaneously binds to the SH3 domain. In cells expressing both receptors, stimulation with androgen or estrogen induces a productive ERα–AR–c-Src complex that can be prevented by treating with either antiandrogens or antiestrogens. PR-B does not bind to c-Src directly, but requires the presence of ERα. In the absence of progestins, PR-B (but not PR-A) binds to ERα. Progestins release ERα from this complex and thus trigger the interaction of ERα with c-Src. In case of ERα, cross-talk is exemplified by epidermal growth factor (EGF) treatment of ovariectomized mice. This results in nuclear translocation and an altered phosphorylation state of ERα. Transient transfection studies showed that EGF, thus, can activate ERα.

The activation of the GR elicits apoptosis in T cells, but can paradoxically also promote thymocyte survival. This dual effect is reminiscent of the actions of the T-cell antigen receptor, which is essential for thymocyte survival but can also cause apoptosis. Interestingly, the activation of the T-cell antigen receptor blocks glucocortiocid-induced apoptosis, thus implying functional cross-talk between these two distinct signaling systems. The regulation of programmed cell death by glucocorticoids also involves the regulated expression and alternative splicing of the *bcl-X* gene and crosstalk with the STAT5 family of transcription factors. In lymphocytes, the apoptotic effect is mediated by a selective interaction of GR with activated STAT5a, whereas in mammary cells, equipped with STAT5b, glucocorticoids act as antiapoptotic factors.

The downstream targets of the rapid action of steroids remain to be fully uncovered, but in many cases the SHRs themselves – and their coactivators – are targeted. Therefore, one function of these rapid actions seems to be the potentiation of genomic action. All SHRs have been known to be phosphoproteins for many years. Most of the identified phosphorylation sites are serine and threonine residues, but some of the family members are also phosphorylated on tyrosine. In case of the chicken PR, for instance, four phosphorylated serines have been identified common to both isoforms, PR-A and PR-B. The two N-terminal sites are only moderately phosphorylated in the absence of hormone, whereas, after hormone treatment, an increase in the phosphorylation of these sites and the appearance of two new phosphorylation sites are observed. The mutation of these serine residues results in a cell- and promoter-specific variation of receptor activity when tested in transfection experiments. In case of ERα ERK, phosphatidylinositol 3-kinase, $p90^{RSK}$, a kinase downstream of Akt2, and Akt2 itself phosporylate discrete residues of the receptor, upregulating its transcriptional activity or stability.

See Also the Following Articles

Adrenal Insufficiency; Androgen Action; Corticosteroids and Stress; Estrogen; Glucocorticoids, Overview; Membrane Glucocorticoid Receptors.

Further Reading

Addison, T. (1855). *On the constitutional and local effects of disease of the suprarenal capsules*. London: D. Highley.

Baulieu, E.-E. and Kelly, P. A. (1990). *Hormones: from molecules to disease*. New York: Chapman and Hall.

Beato, M., Herrlich, P. and Schütz, G. (1995). Steroid hormone receptors: many actors in search of a plot. *Cell* **83**, 851–857.

Cato, A. C. B., Nestl, A. and Mink, S. (2002). Rapid actions of steroid receptors in cellular signaling pathways. *Science's Signal Transduction Knowledge Environment* **138**[Online].

Gronemeyer, H. and Laudet, V. (1995). Transcription factors 3: nuclear receptors. *Protein Profile* **2**, 1173–1308.

Nuclear Receptors Nomenclature Committee. (1999). Nuclear receptors nomenclature committee: a unified nomenclature system for the nuclear receptor superfamily. *Cell* **97**, 161.

Levin, E. R. (2005). Integration of the extranuclear and nuclear actions of estrogen. *Molecular Endocrinology* **19**, 1951–1959.

Lu, N. Z. and Cidlowski, J. A. (2006). Glucocorticoid receptor isoforms generate transcription specificity. *Trends in Cell Biology* **16**, 301–307.

Mangelsdorf, D. J., Thummel, C., Beato, M., et al. (1995). The nuclear receptor superfamily: the second decade. *Cell* **83**, 835–839.

McKenna, N. J. and O'Malley, B. W. (2002). Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell* **108**, 465–474.

Metivier, R., Reid, G. and Gannon, F. (2006). Transcription in four dimensions: nuclear receptor-directed initiation of gene expression. *EMBO Reports* **7**, 161–167.

Parker, M. G. (1991). *Nuclear hormone receptors: molecular mechanisms, cellular functions, clinical abnormalities*. London: Academic Press.

Pratt, W. B., Morishima, Y., Murphy, M., et al. (2005). Chaperoning of glucocorticoid receptors. In: Gaestel, M.

(ed.) *Handbook of experimental pharmacology. Vol. 172: Molecular chaperones in health and disease*, pp. 111–138. Berlin: Springer.
Wu, R. C., Smith, C. L. and O'Malley, B. W. (2005). Transcriptional regulation by steroid receptor coactivator phosphorylation. *Endocrine Reviews* **26**, 393–399.
Wintermantel, T. M., Berger, S., Greiner, E. F., et al. (2005). Evaluation of steroid receptor function by gene targeting in mice. *Journal of Steroid Biochemistry and Molecular Biology* **93**, 107–112.

Relevant Website

NUREBASE, http://www.ens-lyon.fr/LBMC/laudet/nurebase

Steroid Hydroxylases

F H de Jong
Erasmus University Medical Centre, Rotterdam, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F H de Jong, volume 3, pp 505–507, © 2000, Elsevier Inc.

Introduction
Cholesterol Side-Chain Cleavage Enzyme: CYP11A1
17-Hydroxylase: CYP17A1
21-Hydroxylase: CYP21A2
11β-Hydroxylases: CYP11B1 and CYP11B2
Aromatase: CYP19A1
Steroid Catabolism by CYPs

Glossary

Cytochromes P450	Enzymes involved in the hydroxylation of endogenous or exogenous substrates.
Hydroxylation of steroids	Replacement of hydrogen (H) atoms in steroids by hydroxyl (OH) groups.
Localization of steroidogenic enzymes	Specific organs, cell types, or intracellular organelles in which the enzymes, responsible for the biosynthesis of steroids, can be found.
Regulation of steroidogenesis	Processes that play a role in increasing or decreasing the amounts of steroids formed.
Steroid biosynthesis	The process by which hormonally active steroids are formed from cholesterol.

Introduction

The cytochrome P450 (CYP) genes form a family, encoding a class of heme-thiolate enzymes active in the hydroxylation of endogenous and exogenous substances. They are involved in detoxification and catabolism of pharmacological substances and in biosynthesis of steroids. Homology between the various members is relatively high: CYPs in the same subfamily usually are about 70% similar. These enzymes are instrumental in conveying molecular oxygen to designated positions in the steroid skeleton, in this way introducing hydroxyl groups and determining which final hormonal steroids will be produced in the process of steroid biosynthesis. The specific roles of these enzymes in the pathway of steroidogenesis are indicated in **Figure 1.** The electrons necessary to drive this process are supplied by NADPH or NADH and are passed to mitochondrial CYPs by way of the flavoprotein adrenodoxin reductase to the iron-sulfur protein adrenodoxin. For the CYPs in the microsomal cell fraction, the role of adrenodoxin is taken over by P450 oxidoreductase; cytochrome b5 may play an additional role to support 17,20-lyase activity.

Cholesterol Side-Chain Cleavage Enzyme: CYP11A1

The enzyme CYP11A1 performs the conversion of cholesterol to pregnenolone, by first hydroxylating cholesterol at the 20α and 22 carbon atoms and then lysing the bond between these carbon atoms. The enzyme, of which only one form is known, is localized in the adrenal glands, the gonads, the placenta, and the brain. Intracellularly, the location is in the inner mitochondrial membrane. Apparently, this localization is extremely important for the activity of the enzyme: transfection studies with constructs that direct the protein into the endoplasmic reticulum failed to show activity.

Regulation of the expression of CYP11A1 differs in different tissues. In organs where an acute stimulation of CYP11A1 transcription can take place under the influence of adrenocorticotropic hormone (ACTH) or gonadotropins, the adrenals, and the gonads, the responsible cyclic AMP-responsive elements differ from

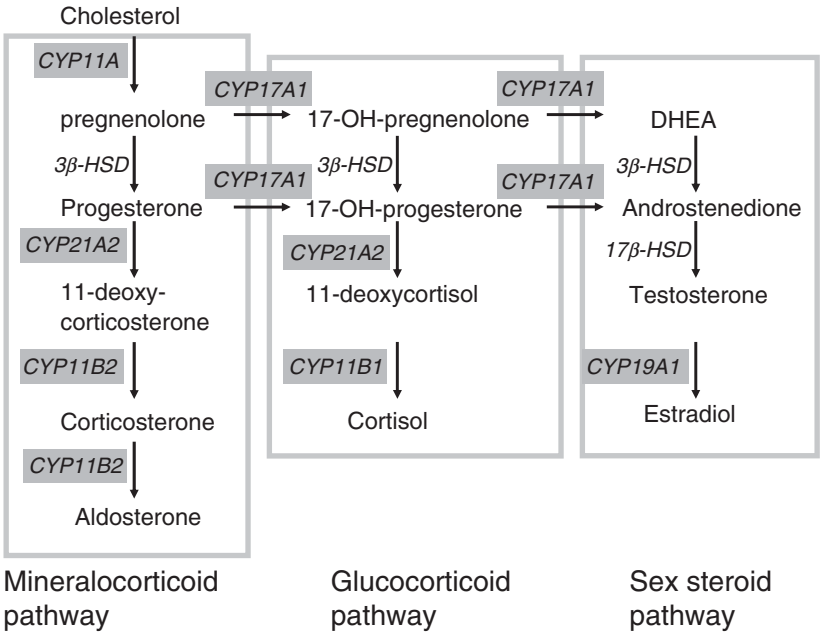


Figure 1 Localization of steroidogenic enzymes along the pathways leading to the biosynthesis of mineralocorticoids, glucocortoids, and sex steroids. Shaded boxes indicate the various cytochromes P450 (CYPs). DHEA, dehydroepiandrosterone.

those in the placenta. Furthermore, a number of transcription factors, such as SF-1 and Sp-1, play a role in the stimulation of the transcription of CYP11A1. This transcriptional regulation of the amount of enzymes in the glands takes a relatively long time (2–4 h), whereas the stimulation of steroidogenesis may be much faster. These fast effects of increased cAMP levels are mediated through effects on the transport of cholesterol into the mitochondria by the steroidogenic acute regulatory protein (StAR), which enables cholesterol transport to the mitochondrial inner membrane. StAR is not expressed in brain or placenta. It has become apparent that the inability to lyse the C20–C22 bond of cholesterol, as present in congenital lipoid adrenal hyperplasia, may be due to inactivating mutations in CYP11A1, but can also be caused by inactivating mutations of StAR, indicating the crucial importance of this protein in steroidogenesis.

17-Hydroxylase: CYP17A1

After the conversion of cholesterol to pregnenolone in the mitochondria, further conversion of pregnenolone takes place in the endoplasmic reticulum by the conversion to 17-OH pregnenolone by CYP17A1 or to progesterone by 3β-HSD. Progesterone, in turn, can be converted to 17-OH progesterone by the same CYP17A1, which also contains 17,20-lyase activity to convert 17-OH pregnenolone to dehydroepiandrosterone (DHEA) and 17-OH progesterone to androstenedione. For this reason, CYP17A1 is at the bifurcation of the steroidogenic pathways to glucocorticosteroids and the sex steroids. CYP17 is expressed in the adrenal glands and in the gonads. The production of DHEA, which might be important as a neurosteroid through its interactions with γ-aminobutyric acid (GABA) and N-methyl-D-aspartate (NMDA) receptors, in the brain is an enigma so far because CYP17A1 has not been detected in the brain, which has led to alternative hypotheses for the mechanisms responsible for the production of this steroid involving reactive oxygen species. The expression of the CYP17A1 gene is mainly regulated by cAMP. Addition of cAMP derivatives can stimulate the transcription of the gene; this activity can be inhibited by the addition of phorbol esters, which act via protein kinase C and growth factors such as epidermal growth factor (EGF) and fibroblast growth factor (FGF).

The separate regulation of 17-hydroxylase and 17,20-lyase activity of the enzyme has been resolved on the basis of the finding of mutated genes, causing the production of 17-hydroxylated steroids, without removal of the C20,C21 side chain. These mutations are localized in the region of the protein, which interacts with the additional electron donor cytochrome b5; lyase activity appears to be more vulnerable to the disruption of interaction with this redox partner than the 17-hydroxylase reaction. Similarly, the changing ratios between the production of glucocorticoids and adrenal androgens, mainly DHEA

sulfate, during human development is likely to be due to a changed expression of cytochrome b5.

The pivotal role of CYP17A1 in steroidogenesis is emphasized by the effects of inactivating mutations in this enzyme: a complete block of 17-hydroxylation leads to the inability to synthesize cortisol, whereas aldosterone will be produced in abundance, giving rise to hypertension. Furthermore, sex steroids cannot be produced, causing undervirilization in boys and absence of puberty in both sexes.

21-Hydroxylase: CYP21A2

CYP21A2 catalyzes the conversion of progesterone to 21-hydroxyprogesterone (deoxycorticosterone, DOC) and of 17-OH progesterone to 17,21-dihydroxyprogesterone (deoxycortisol, compound S). These reactions are a prerequisite to synthesize mineralocorticoids and glucocorticoids; the enzyme is expressed in the cortex of the adrenal gland and is not present in the gonads. Using reverse transcription polymerase chain reaction (RT-PCR), expression of CYP21A2 has also been detected in the brain stem, but not in other regions of the brain.

The expression of CYP21A2 is stimulated by cAMP- and protein kinase C-dependent pathways. In this way, the production of deoxycorticosterone in the zona glomerulosa of the adrenal cortex may be stimulated by angiotensin II, whereas the production of deoxycortisol in the zonae fasciculata and reticularis is ACTH dependent.

Due to the fact that an inactive pseudogene, CYP21A1P, exists in the neighborhood of the CYP21A2 gene on chromosome 6, frequent exchanges between these genes have taken place, leading to impairment of the functioning of the resulting protein, the inability to produce aldosterone, corticosterone, and cortisol, and overproduction of adrenal androgens (in humans). These enzyme deficiencies can lead to salt loss in newborn babies and severe masculinization in girls. In female patients with milder forms of the defect, late onset virilization or hirsutism may be found.

11β-Hydroxylases: CYP11B1 and CYP11B2

The further conversion of 21-hydroxylated steroids to the main mineralocorticoids and glucocorticoids (aldosterone and cortisol [humans] or corticosterone [rat and mouse]) is affected by CYP11B2 and CYP11B1, respectively. The genes encoding these members of the CYP family are found in close vicinity on chromosome 8q and probably arise from ancient gene duplication. The localization of these enzymes in the adrenal cortex is different: CYP11B1 is localized in mitochondria of the zona fasciculata, whereas CYP11B2 is found in mitochondria of the zona glomerulosa. Furthermore, using RT-PCR, mRNAs encoding these enzymes have been detected in various regions of the brain and in the heart.

The regulation of the expression of the two genes is also different. The expression of CYP11B1 is regulated by the ACTH-driven production of cAMP, followed by activation of a cAMP-responsive element, whereas CYP11B2 is stimulated mainly by angiotensin II, which stimulates the production of inositol triphosphate and diacylglycerol by phospholipase C. This leads to increased levels of intracellular calcium and increased activity of phosphokinase C. In this way, the expression of the two 11-hydroxylases can be regulated separately in the zona glomerulosa and zona fasciculata, respectively, of the adrenal cortex.

The clinical results of deficiencies of either of these enzymes can be predicted on the basis of the steroids, the biosynthesis of which will be blocked. Defective CYP11B1 gives rise to low levels of cortisol that can only be increased by supply of exogenous glucocorticoids. An insufficient production of aldosterone, caused by a mutation in CYP11B2, can be partially compensated by overproduction of deoxycorticosterone, which has mineralocorticoid activity.

Aromatase: CYP19A1

The final steroid-hydroxylating enzyme, which plays a role in the biosynthesis of steroids, is CYP19A1. It is responsible for the hydroxylation of carbon atom C19 of testosterone and androstenedione, the subsequent cleavage of the bond between C19 and C10, and subsequent redistribution of the electrons over the A ring to form the aromatic steroids estradiol and estrone, respectively. Only one copy of the gene has been found. It has a complex structure at the 5′ end, allowing start of transcription at a number of sites, under the control of various promoter sequences. This means that in various tissues, different mechanisms can be used for the stimulation of the transcription of the gene: in the gonads and adrenals, a cAMP-mediated mechanism is responsible for the stimulation of CYP19A1 by gonadotropins and ACTH, respectively, whereas in fat cells, cortisol can stimulate the production of estrogens. CYP19A1 is also expressed in a large number of other tissues, such as muscle, bone, and brain. In the latter, estrogens formed by this enzyme may play a role in sexual differentiation.

Defects of CYP19A1 are very rare. However, the clinical picture in patients with mutations in this gene shows profound changes of parameters of gonadal function, bone metabolism, and aspects of the

metabolic syndrome, indicating the role of estrogens in the regulation of numerous aspects of homeostasis.

Steroid Catabolism by CYPs

Apart from their role in the biosynthesis of steroid hormones, enzymes of the CYP family also play important roles in the breakdown of steroids. The most important enzymes in this respect are those of the CYP1 and CYP4 subfamilies. Members of the CYP1 subfamily are important in the hydroxylation of mainly estrogens, where both C2 and C4 may be attacked. CYP4 family members hydroxylate corticosteroids and androgens at the 6 position and can also attach a hydroxyl group at C16. It is important to be aware of the fact that expression of all of these enzymes can be induced by xenobiotics, leading to aberrant metabolism of the steroid hormones and, for example, changes in the ratios between the serum levels of steroids produced in the adrenal gland, which might lead to erroneous conclusions of disturbed steroidogenesis.

See Also the Following Articles

11β-Hydroxysteroid Dehydrogenases; Estrogen; Steroid Hormone Receptors.

Further Reading

de Ronde, W., Pols, H. A. P., van Leeuwen, J. P. T. M., et al. (2003). The importance of oestrogens in men. *Clinical Endocrinology* **58**, 529–542.

Fujieda, K., Okuhara, K., Abe, S., et al. (2003). Molecular pathogenesis of lipoid adrenal hyperplasia and adrenal hypoplasia congenita. *Journal of Steroid Biochemistry and Molecular Biolology* **85**, 483–489.

Grinberg, A. V., Hannemann, F., Schiffler, B., et al. (2000). Adrenodoxin: structure, stability, and electron transfer properties. *Proteins* **40**, 590–612.

Lewis, D. F. (2004). 57 varieties: the human cytochromes P450. *Pharmacogenomics* **5**, 305–318.

Lisurek, M. and Bernhardt, R. (2004). Modulation of aldosterone and cortisol synthesis on the molecular level. *Molecular and Cellular Endocrinology* **215**, 149–159.

Manna, P. R., Wang, X. J. and Stocco, D. M. (2003). Involvement of multiple transcription factors in the regulation of steroidogenic acute regulatory protein gene expression. *Steroids* **68**, 1125–1134.

Nebert, D. W. and Russell, D. W. (2002). Clinical importance of the cytochromes P450. *Lancet* **360**, 1155–1162.

New, M. I. (2003). Inborn errors of adrenal steroidogenesis. *Molecular and Cellular Endocrinology* **211**, 75–83.

Shimada, T. and Fujii-Kuriyama, Y. (2004). Metabolic activation of polycyclic aromatic hydrocarbons to carcinogens by cytochromes P450 1A1 and 1B1. *Cancer Science* **95**, 1–6.

Simpson, E. R., Misso, M., Hewitt, K. N., et al. (2005). Estrogen—the good, the bad, and the unexpected. *Endocrine Reviews* **26**, 322–330.

Stratakis, C. A. and Bossis, I. (2004). Genetics of the adrenal gland. *Reviews in Endocrine and Metabolic Disorders* **5**, 53–68.

You, L. (2004). Steroid hormone biotransformation and xenobiotic induction of hepatic steroid metabolizing enzymes. *Chemico-Biological Interactions* **147**, 233–246.

Steroid Synthetic Pathways *See:* Steroid Hydroxylases; 11β-Hydroxysteroid Dehydrogenases.

Steroidogenesis Inhibitors *See:* 11β-Hydroxysteroid Dehydrogenases.

Strain Differences in Stress Response in Rodents

P Mormède
INRA – Université Victor Segalen, Bordeaux, France

© 2007 Elsevier Inc. All rights reserved.

Strain Differences
Sources of Genetic Variability
Systems Genetics and Genetic Reference Populations

Glossary

Inbred	Obtained by continuous breeding of closely related individuals so as to preserve desirable traits in a stock.
Inbred strain	A strain of organisms that has been maintained by sibling (sister × brother) mating for 20 or more consecutive generations. Except for the sex differences, mice of an inbred strain are as genetically alike as possible, being homozygous at virtually all of their loci.
Isogenic	Genetically identical (except for sex); coming from the same individual or from the same inbred strain.
Intercross	A cross between two animals that have the same heterozygous genotype at designated loci; for example, between sibling F1 hybrids that were derived from an outcross between two inbred strains.
Quantitative trait	A phenotype that is influenced by multiple genes. Its variation is therefore continuous (skin color, weight, etc.).
Quantitative trait locus (QTL)	A polymorphic site on a chromosome containing alleles that differentially influence the expression of a quantitative trait.
Recombinant inbred (RI) strains	Strains constructed by crossing two inbred strains to produce an F1 generation, followed by 20 or more consecutive generations of brother × sister mating. However, the F1 generation may be produced in other ways, such as crossing mice of either existing RI strains, advanced intercross lines, or multiple inbred strains (e.g., an eight-way cross).
Systems genetics	A method for integrating across scales of organization using genetic reference populations. It summarizes results on molecular networks (gene expression data) associated with single-nucleotide polymorphisms SNPs and structural and functional variations.

Strain Differences

Inbred strains of mice and rats are invaluable animal resources for studying the genetic bases of phenotypic variation. They have been obtained through brother–sister mating for more than 20 generations, so all individuals have the same genome (isogenic) and each individual is homozygous at each locus. In a given environment, strain differences result from differences at the genome level.

Large variations in hypothalamic-pituitary-adrenal (HPA) axis activity have been described among inbred strains of rats and mice. Many of these studies are related to the pathophysiology of immune diseases, in which the HPA axis could play a critical role together with immune mechanisms. The Fischer 344 (F344) and Lewis rat strains particularly were studied in this respect, the Lewis female rats bearing an HPA axis that is less reactive to most stimuli and being more sensitive to induced autoimmune diseases such as streptococcal cell wall-induced arthritis or experimental allergic encephalitis. The sensitivity of the Lewis strain can be restored by the removal of the adrenal glands or by treatment with a glucocorticoid antagonist, and F344 rats become resistant when treated with glucocorticoid hormones. They are also studied in relation with the metabolic effects of corticosteroid hormones in the context of nutrition, obesity, and cardiovascular diseases.

Sources of Genetic Variability

It is not easy to detect the primary change(s) responsible for the functional variations of the HPA axis because phenotypic differences may be multiple and the components of the HPA axis interact strongly with numerous facilitation and feedback mechanisms at every level. For instance, when compared to F344 rats, the inbred brown Norway (BN) rats display a number of differences in their HPA axis function, with a different diurnal pattern of plasma corticosterone levels, a faster recovery after restraint stress, an adrenal gland of larger size but less reactive to adrenocorticotropic hormone (ACTH), and mineralocorticoid receptors (MRs) and glucocorticoid receptors (GRs) hypersensitive to their ligands.

Two main strategies are available for exploring the molecular bases of genetic variation. The detailed exploration of phenotypic differences may suggest candidate mechanisms and/or genes to analyze. For

instance, the comparison of BN and F344 rat strains suggested that the MR is a major source of genetic variability. Therefore, the gene encoding MR (named Nr3c2) was cloned in the two strains and sequenced, revealing in the BN strain a mutation in the coding region with an amino acid change in the protein (Y73C) that increases the transactivating properties of MR agonists.

Another approach, known as quantitative trait loci (QTL) analysis, is based on the association between polymorphic anonymous markers and the phenotypical value of the trait under study in a segregating population (such as F2 or backcross). In this approach, there is no need for any assumption about the gene(s) involved. In the case of the BN/F344 comparison, a QTL analysis confirmed the linkage between the mutation in the Nr3c2 gene and several MR-related phenotypes and disclosed several other loci contributing to phenotypic variability of MR- and GR-related phenotypes. Another study in a BN × Wistar Kyoto hyperactive (WKHA) intercross revealed the contribution of a locus on chromosome 2 influencing poststress levels of renin and aldosterone. In this locus the angiotensin receptor 1b was mapped, which is responsible for angiotensin II (Ang II) effects on aldosterone secretion, and, indeed, differences were found in the response of adrenocortical cell to Ang II stimulation *in vitro*. This example illustrates the power of the QTL approach to reveal candidate genes possibly involved in phenotypic variation, as well as the multiplicity of mechanisms underlying genetic variation in corticosteroid hormone secretion and action.

Systems Genetics and Genetic Reference Populations

An extension of this QTL approach is the use of recombinant inbred (RI) strains, which are constructed by crossing two inbred strains to produce an F1 generation, followed by 20 or more consecutive generations of brother–sister mating. From a single F1 population, as many strains as possible can be derived. Each strain contains a unique, approximately equal proportion of genetic contributions from the two progenitor inbred strains. Instead of having a single individual for each genetic combination, as in an F2 population, each strain represents an unlimited number of animals with the same genotype, so the phenotype can be measured with more accuracy (in several animals instead of a single F2 animal) and different phenotypes measured across studies are cumulative and can be compared. Furthermore, precise genotyping of the strains with sequence-derived single nucleotide polymorphisms allows precise QTL mapping. For instance the most used set of recombinant inbred strains with approximately 80 strains has been obtained from the C57BL/6 and DBA/2 mouse strains, and much larger sets (up to 1000 strains) are under construction. All data about these strains are integrated in a unique database named WebQTL that allows a large range of data mining from genetic correlation across multiple traits and QTL analysis. As an example, plasma corticosterone levels after an injection of saline (which can be considered as a mild stressor) was studied in 19 strains. In the WebQTL database of published phenotypes, genetic correlations are found between these levels and a number of phenotypes related to neuroanatomical and neurofunctional measures, response to drugs of abuse, and immune and inflammatory processes. This can be also correlated with gene expression data in different tissues. Therefore, these genetic reference populations allow a systems genetics approach that integrates all these data from diverse sources to explore genetic variation in HPA axis and in its target organs as a complex functional system.

See Also the Following Articles

Corticosteroid Receptors; Hypothalamic-Pituitary-Adrenal; Genetic Variation of HPA Axis Activity and Function in Farm Animals; Genetic Testing and Stress.

Further Reading

Chesler, E. J., Lu, L., Wang, J., et al. (2004). WebQTL: rapid exploratory analysis of gene expression and genetic networks for brain and behavior. *Nature Neuroscience* **7**, 485–486.

Churchill, G. A., Airey, D. C., Allayee, H., et al. (2004). The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nature Genetics* **36**, 1133–1137.

Marissal-Arvy, N., Lombes, M., Petterson, J., et al. (2004). Gain of function mutation in the mineralocorticoid receptor of the Brown Norway rat. *Journal of Biological Chemistry* **279**, 39232–39239.

Mormède, P., Courvoisier, H., Ramos, A., et al. (2002). Molecular genetic approaches to investigate individual variations in behavioral and neuroendocrine stress responses. *Psychoneuroendocrinology* **27**, 563–583.

Peirce, J. L., Lu, L., Gu, J., et al. (2004). A new set of BXD recombinant inbred lines from advanced intercross populations in mice. *BMC Genetics* **5**, 7–8.

Relevant Websites

Jackson Laboratory. www.jax.org.
WebQTL. www.genenetwork.org.

Stress and Anxiety: Treatment with 2nd Generation Antipsychotic Drugs

J S Ballon
Stansford University, Stansford, CA, USA
J A Boyd and D A Wirshing
VA Greater Los Angeles Healthcare System, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Second-Generation Antipsychotic Drugs as Anxiolytic Agents in Animal Models of Anxiety
Akathisia – More Prevalent than it Appears?
The 5HT1A Receptor – an Anxiety Receptor for Antipsychotic Drugs?
Histamine – Antipsychotic Drugs as Soporific Agents
Cortisol – How do Stress Hormones Respond to Dopamine Blockade?
Clinical Reports of Second-Generation Antipsychotic Drugs in Anxiety Disorders
Future Directions

Glossary

Akathisia	A state of inner restlessness characterized by pacing, inability to sit still, or other constant urge for movement.
Antipsychotic	A medication used to treat psychosis, i.e. delusions or hallucinations.
Anxiolytic	A medication or treatment used to decrease anxiety.
Anxiogenic	A factor which causes or increases levels of anxiety.
Cholinergic	Related to the neurotransmitter acetylcholine.
Dopaminergic	Related to the neurotansmitter dopamine.
GAD	Generalized anxiety disorder.
Histaminergic	Related to the neurotransmitter histamine.
Hypercortisolemia	An elevated level of the glucocorticoid cortisol in the blood.
OCD	Obsessive-compulsive disorder.
Presynaptic	A receptor located before the connection of nerves (synapse) which influences the nature of the synapse.
SANS	Scale for the assessment of negative symptoms.
Serotonergic	Related to the neurotransmitter serotonin.

Second-generation antipsychotic medications are a heterogeneous group of medications that work at varying levels at many receptors, including dopaminergic, serotonergic, cholinergic, and histaminergic. Varying potencies at these receptors, and specific receptor subtypes, are responsible for not just the antipsychotic and antimanic properties that these medications all share, but may also account for varying levels of anxiolytic effect. This article will explore some mechanistic answers for this anxiolytic effect including the effect on cortisol, decreased levels of akathisia compared to first-generation antipsychotic drugs, histaminergic effects, and increased potency at the serotonergic (5HT) 1A receptor.

Second-Generation Antipsychotic Drugs as Anxiolytic Agents in Animal Models of Anxiety

Initial efficacy for anxiolytic properties for antipsychotic drugs was controversial, but first observed in rat data. Bruhwyler et al. demonstrated that rats in an open-field paradigm had greater amount of exploration and decreased defecation while on high doses of clozapine. This was in contrast to the effect of haloperidol in a similar paradigm. Additionally, in an elevated plus-maze protocol, rats on a low dose of clozapine demonstrated behavior consistent with anxiogenic properties as they spent less time in the open part of the maze compared with rats who received only the vehicle injection.

Inoue et al. demonstrated that olanzapine and clozapine worked to prevent the acquisition of freezing, but not the expression of freezing, in a conditioned-fear paradigm. This was contrasted however with benzodiazepines, selective serotonin-reuptake inhibitors (SSRIs) and 5HT1A antagonists which both inhibit the acquisition and expression of freezing. However, there was a slight separation from haloperidol and chlorpromazine in the response to the conditioned stimulus.

Further evidence of anxiolytic properties of second-generation antipsychotic drugs was shown in a conflict schedule model in which rats on olanzapine demonstrated similar behavior to those on chlordiazepoxide, but those on haloperidol and risperidone did not. In contrast, risperidone was shown to have short-term anxiolytic properties in a study by Nowakowska et al., however these effects were not seen after 7 days of treatment.

Akathisia – More Prevalent than it Appears?

Akathisia is a troubling side effect that is seen in upwards of 90% of younger patients treated with

first-generation antipsychotic drugs. In a mixed inpatient population, the rate of akathisia in patients on first-generation antipsychotic medications was 18% and pseudoakathisia was 24%. Though the prevalence tends to decrease with age, even in patients aged over 65 years, it is seen in up to 15% of patients.

These data are especially problematic as the core symptoms of akathisia are severe internal restlessness which generally leads to increased motor activity including pacing, marching, and ultimately can manifest as a severe dysphoria. Other common symptoms of akathisia include tension, panic, irritability, and impatience. Suicidality has been linked to akathisia in patients on high potency first-generation antipsychotic drugs. Akathisia is difficult for patients to describe subjectively and can, therefore, often be confused for anxiety. Patients who develop akathisia are also at a much higher risk for nonadherence to treatment and therefore are at a higher risk for relapse.

Second-generation antipsychotic drugs have shown a decreased rate of akathisia. This is considered to be related to the presynaptic blockade of 5HT2A receptors which lead to increased levels of dopamine in the nigrostriatal pathway compared to first-generation antipsychotic drugs. In several trials, olanzapine has been shown to have decreased rates of extrapyramidal side effects including akathisia. In a large meta-analysis, olanzapine, risperidone, and placebo were shown to have decreased rates of akathisia compared with haloperidol.

In addition to the decreased rates of akathisia seen with second-generation antipsychotic drugs, there has been a benefit shown in treatment emergent extrapyramidal symptoms, including akathisia, when patients are switched from haloperidol to olanzapine. Costa e Silva et al. demonstrated an 83% decrease in akathisia symptoms, as measured by the Barnes Akathisia Scale, in patients switched from haloperidol to olanzapine in a large, multicenter trial in Latin America. It may be that some of the clinical observations of anxiety in patients taking first-generation antipsychotic medications is related to under diagnosed akathisia and that anxiolytic effects of second-generation antipsychotic drugs are related in part to the decreased rate of akathisia seen.

The 5HT1A Receptor – an Anxiety Receptor for Antipsychotic Drugs?

The 5HT1A receptor has long been associated with anxiolytic benefit. Buspirone is a partial agonist at 5HT1A receptor and has been considered by many to be an anxioselective agent. Atypical antipsychotic medications have varying degrees of activity at the 5HT1A receptor, with the strongest affinity seen by aripiprazole and ziprasidone. 5HT1A agonists do not typically occupy the receptor at clinically significant doses, but do alter the overall activity at the receptor. Aripiprazole, ziprasidone, quetiapine, and clozapine have been shown to have partial agonist activity at 5hT1A, with aripiprazole and ziprasidone the most potent. 5HT1A agonism has been linked to increased release of dopamine in the striatum, though it is not clear if this is clinically significant.

Ziprasidone has shown anxiolytic effects in non-psychotic individuals. In a double-blind, randomized trial (30 subjects), ziprasidone was compared to diazepam for anxiety prior to dental procedures. Ziprasidone showed a more rapid, and a slightly more potent, anxiolytic effect compared to diazepam which had a later onset and duration of action. Both were effective, with ziprasidone showing a lesser effect on blood pressure and not showing any greater sedation than placebo, compared to diazepam which was substantially sedating for over 2 hours.

Histamine – Antipsychotic Drugs as Soporific Agents

Histaminergic agents have long been used as treatments for anxiety and insomnia. Histamine antagonists, such as hydroxyzine, have been used as successful treatments in generalized anxiety disorder (GAD), and are often used in posttraumatic stress disorder (PTSD) and acute anxiety. Olanzapine and quetiapine are potent antagonists of the H1 receptor. H1 antagonists have also been shown to reduce the level of cortisol, which may be an indirect mechanism for achieving an anxiolytic effect. In this way, the antihistaminergic antipsychotic drugs share characteristics with some of the tricyclic antidepressant drugs, such as clomipramine and imipramine. Clomipramine, a highly histaminergic tricyclic antidepressant, remains one of the most powerful medications for obsessive–compulsive disorder (OCD) currently available. Recent work has identified antipsychotic drugs as useful in augmentation of OCD treatment. Risperidone has the most robust data, but there are reports regarding the use of quetiapine and olanzapine. Antipsychotic drugs have not shown benefit as monotherapy however, but in combination with serotonergic agents provide additional benefit in treatment refractory cases. The data for patients with comorbid tic disorders shows the most benefit from treatment with antipsychotic drugs.

Antipsychotic drugs are also being increasingly prescribed for insomnia. In patients with insomnia secondary to PTSD, olanzapine has shown a benefit as

an augmenting agent, targeting sleep and nightmares. There have been some initial studies regarding alpha-1 antagonists such as prazosin in decreasing nightmares associated with PTSD. Olanzapine and quetiapine are potent alpha-1 antagonists as well, and in that way may provide a benefit in treatment of nightmares and insomnia associated with PTSD. Additionally, the antihistamine properties of these medications also make them sedating and while that may be a treatment limiting side effect in certain circumstances, makes them beneficial agents in the treatment of sleep. Despite these benefits, caution should be exercised in using these medications solely for treatment of insomnia as there are many other treatment options and the side effect profile of these medications restricts their use to only the most refractory cases in which polypharmacy is needed.

Cortisol – How do Stress Hormones Respond to Dopamine Blockade?

Cortisol is a hormone that is typically associated as a biomarker for stress that has also been shown to be elevated in depression and anxiety. It is also been shown to be elevated overall in schizophrenia. A correlation has been seen between increased negative symptoms, measured by the Scale for the Assessment of Negative Symptoms (SANS), and elevations in cortisol, though overall psychopathology scores, medication dosage, and type of medication needed do not necessarily correlate with cortisol levels. Cognitive symptoms, often considered to be most predictive of future prognosis in schizophrenia, are also correlate severity with increased cortisol levels.

The worsened cognitive symptoms with hypercortisolemia may be related to long-term damage to the hippocampus that is seen with chronic elevations in cortisol. In the primate brain, high concentrations of glucocorticoid receptors are seen in the hippocampus and in the primates given chronic dexamethasone treatment, there is an approximately 30% reduction in size and segmental volumes of the hippocampus over time. Similar effects on the hippocampus are seen in primates exposed to animal paradigms of chronic stress. Schizophrenia patients, even at initial presentation, are known to have structural deficits of the hippocampus.

In a double-blind, randomized, cross-over trial, olanzapine and quetiapine showed decreased cortisol levels compared to haloperidol and placebo. The mechanism behind such a direct action on cortisol is not fully understood, but may be most prominently related to the action on serotonergic receptors, with some contribution from histaminergic or adrenergic blockade. The decreased level of cortisol may relate to the antidepressant and anxiolytic action seen with these agents over the first-generation antipsychotic drugs.

Clinical Reports of Second-Generation Antipsychotic Drugs in Anxiety Disorders

The clinical literature on using antipsychotic medications for anxiety disorders is fairly limited. The mainstay of treatment are primarily the SSRIs and serotonin-norepinephrine reuptake inhibitors (SNRIs) and thus antipsychotic medications are typically used adjunctively when those medications are not fully effective. Other medications that are also utilized include benzodiazepines and antihistaminergic medications and buspirone. Atypical antipsychotic medications are overall used mostly as adjunctive agents in the treatment of refractory anxiety.

A double-blind, placebo-controlled trial of risperidone augmentation in GAD showed decreased core anxiety symptoms on the Hamilton Rating Scale for Anxiety. There have been several small, open label studies in GAD and social anxiety disorder that have had conflicting and ultimately inconclusive results. Risperidone, ziprasidone, olanzapine, quetiapine, and aripiprazole all have open-label studies associated with their use in refractory GAD, but thus far none have been demonstrated as effective monotherapy. While the data is scant, what can be surmised is that there is a subset of patients with severe GAD who are aided by these medications. Particularly, the medications that are most soporific will induce sleep and can be calming to the patients suffering from GAD. Further work and double-blind studies are needed to determine the ultimate clinical benefit from these agents.

In PTSD, atypical agents have most been targeted for their effects on sleep and insomnia. Nightmares and other intrusive thoughts are a common feature of this disorder and have been shown to be generally responsive to atypical antipsychotic drugs. In war survivors in Croatia, monotherapy with fluphenazine, risperidone, quetiapine, or olanzapine led to decreased symptoms of PTSD. The more sedating antipsychotic drugs have shown an increased benefit with promoting more restful sleep with the most data being shown for quetiapine. Other data have been shown in olanzapine and risperidone. Again, these agents have shown the most promise when used adjunctively to serotonergic antidepressants.

There are a number of articles in which antipsychotic drugs are used to control severe OCD symptoms, however, the paradox is that there are also a

number of articles suggesting that the atypical antipsychotic medications may worsen obsessive compulsive symptoms and even bring them on *de novo*. The thought is that the 5HT2A antagonism is the mechanism by which the anxiety/OCD symptoms can worsen.

Future Directions

There is growing support and evidence that antipsychotic drugs may provide anxiolytic benefit for a variety of disorders – psychotic disorders as well as PTSD, OCD, and GAD. The future challenges we face in using this class of agents so broadly is to do so carefully, as there are many untoward side effects that are emerging, most concerning are diabetes, weight gain, and dyslipidemia. Chronic usage of atypical antipsychotic drugs requires close health maintenance and monitoring for subtle changes in metabolic activity. More rare side effects, such as neuroleptic malignant syndrome and dystonia, must also not be forgotten, as they are reported in all the atypical antipsychotic medications, though at frequencies that are less than the first-generation, high-potency agents. As the treatment of anxiety disorders has shifted more towards the primary care setting, education into the less common side effects is important in those prescribing these agents. Additional consideration must be made towards the relatively high cost of these medications compared to other alternatives. Atypical antipsychotic drugs do not characteristically show addictive potential, though there have been reports of diversion of quetiapine, particularly in the prison population. Overall, these agents do show promise, particularly for the refractory patient, and when used judiciously provide a helpful adjunct to therapy for severe anxiety disorders.

Further Reading

Ahearn, E. P., Winston, E., Mussey, M. and Howell, T. (2003). Atypical antipsychotics, improved intrusive symptoms in patients with posttraumatic stress disorder. *Military Medicine* **168**(9), x.

Baker, R. W., Ames, D., Umbricht, D. S., Chengappa, K. N. and Schooler, N. R. (1996). Obsessive-compulsive symptoms in schizophrenia: a comparison of olanzapine and placebo. *Psychopharmacology Bulletin* **32**(1), 89.

Bartzokis, G., Lu, P. H., Turner, J., Mintz, J. and Saunders, C. S. (2005). Adjunctive risperidone in the treatment of chronic combat-related posttraumatic stress disorder. *Biological Psychiatry* **57**(5), 474.

Brawman-Mintzer, O., Knapp, R. G. and Nietert, P. J. (2005). Adjunctive risperidone in generalized anxiety disorder: a double-blind, placebo-controlled study. *Journal of Clinical Psychiatry* **66**(10), 1321.

Cohrs, S., Röher, C., Jordan, W., et al. (2006). The atypical antipsychotics olanzapine and quetiapine, but not haloperidol, reduce ACTH and cortisol secretion in healthy subjects. *Psychopharmacology* **185**(1), 11.

Inoue, T., Tsuchiya, K. and Koyama, T. (1996). Effects of typical and atypical antipsychotic drugs on freezing behavior induced by conditioned fear. *Pharmacology, Biochemistry and Behavior* **55**(2), 195.

Nemeroff, C. B. (2005). Use of atypical antipsychotics in refractory depression and anxiety. *The Journal of Clinical Psychiatry* **66**(Suppl 8), 13.

Nowakowska, E., Chodera, A., Kus, K. and Rybakowski, J. (1999). Some behavioural effects of risperidone in rats: comparison with haloperidol. *European Neuropsychopharmacology* **9**(5), 421.

Parker, K. J., Schatzberg, A. F. and Lyons, D. M. (2003). Neuroendocrine aspects of hypercortisolism in major depression. *Hormones and Behavior* **43**(1), 60.

Pierre, J. M., Shnayder, I., Wirshing, D. A. and Wirshing, W. C. (2004). Intranasal quetiapine abuse. *American Journal of Psychiatry* **161**(9), 1718.

Pollack, M. H., Simon, N. M., Zalta, A. K., et al. (2006). Olanzapine augmentation of fluoxetine for refractory generalized anxiety disorder: a placebo controlled study. *Biological Psychiatry* **59**(3), 211.

van Liempt, S., Vermetten, E., Geuze, E. and Westenberg, H. G. M. (2006). Pharmacotherapy for disordered sleep in post-traumatic stress disorder: a systematic review. *International Clinical Psychopharmacology* **21**(4), 193.

Wirshing, D. A. (2004). Schizophrenia and obesity: impact of antipsychotic medications. *The Journal of Clinical Psychiatry* **65**(Suppl 18), 13.

Wirshing, D. A., Boyd, J. A., Meng, L. R., Ballon, J. S., Marder, S. R. and Wirshing, W. C. (2002). The effects of novel antipsychotics on glucose and lipid levels. *Journal of Clinical Psychiatry* **63**(10), 856.

Wirshing, W. C. (2001). Movement disorders associated with neuroleptic treatment. *Journal of Clinical Psychiatry* **62**(Suppl 21), 15.

Stress and CNS Arousal: Genomic Contributions

A C Ribeiro and D W Pfaff
Rockefeller University, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

Arousal
Neuroanatomical and Neurochemical Basis of Arousal
Amplified States of Arousal
Significance of Studying Arousal
Promising Future Lies Ahead in Arousal Research

Glossary

Arousal	The fundamental property of the central nervous system that determines the amplitude of responses to any given stimulus.
Arousal crescent	A group of primitive neurons residing in the ventral and medial edges of the brain stem that have ascending projections that alert the brain.
Arousal-promoting mechanisms	Monoaminergic cell groups (dorsal raphe nuclei, locus coeruleus, substantia nigra and ventral periaqueductal gray, and tuberomammillary nuclei) that have ascending projections to the thalamus, hypothalamus, and cerebral cortex and whose chemical products favor states of behavioral arousal.
Sleep-promoting mechanisms	A small group of neurons residing in the ventrolateral preoptic nucleus that fire predominantly during sleep and send inhibitory projections to the monoaminergic cell groups.

Arousal

Stress responses, whether defined endocrinologically or behaviorally, depend for their expression on the fundamental arousal of the central nervous system (CNS). The relation between stress and arousal is not simple. In terms of a person's psychological feelings, moderate and increasing levels of arousal can be a good thing. However, too much or for too long is both tiring and unpleasant (Figure 1).

In the workplace, the relationship between stress and arousal has been defined by occupational psychologists (Table 1). Too much demand with too little control can be so stressful as to cause illness. Therefore, it is useful to explore CNS mechanisms underlying the concept of arousal. Here we briefly summarize its operational definition, neuroanatomy, functional genomics, and neurochemistry.

Generalized Arousal

Although intuitively simple to understand, the concept of arousal has not been an easy one to define scientifically. Because all our behaviors and responses to stimuli depend on the underlying state of arousal we are in, it becomes critically important to quantitatively define arousal and to have ways to accurately measure it.

Arousal refers to the degree of stimulation necessary to drive a certain behavior in an animal. The particular stimulus or behavior can vary, but the underlying arousal state determines the magnitude of the response and even if the behavior will occur at all. Another way of envisioning this concept is that arousal is the force or motivation behind a given response, whereas the stimulus dictates what the response modality is.

The overall arousal state of an organism depends on generalized arousal and many specific forms of arousal, such as hunger, thirst, sex, sleep, and fear (Figure 2). If we express this concept in the framework of a mathematical equation, then arousal (A) is a function of the general arousal (A_g) state and each individual specific arousal component (A_{s1-n}). Influencing each of these variables is the underlying temperament of the individual, which we express as a constant (K) for each variable (g, $1–n$). Because these parameters (thirst, hunger, sex, sleep, fear, etc.)

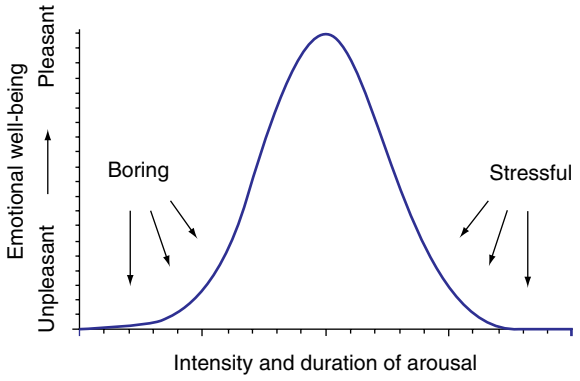


Figure 1 Prolonged arousal can lead to stress.

Table 1 Relationship between occupational demands and employee's control

		Demand for arousal, alertness, and vigilance	
		High	*Low*
Degree of control by employee	High	Stimulating	Boring
	Low	Stressful	Restful

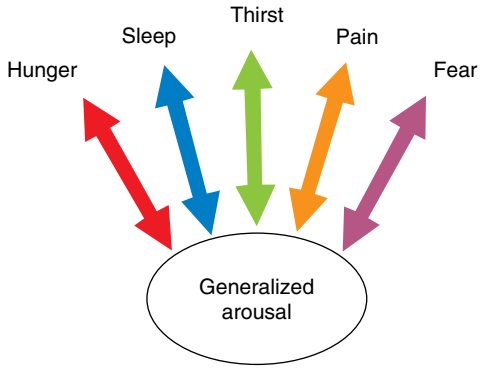


Figure 2 Representation of the arousal equation. Each specific arousal (e.g., pain) depends on the underlying generalized arousal state of the animals, as well as on other specific arousal components (e.g., hunger motivation).

often influence one another, the Arousal Equation can quickly become very complex:

$$A = f(\mathrm{K}_g A_g + \mathrm{K}_{s1} A_{s1} + \mathrm{K}_{s2} A_{s2} + \ldots + \mathrm{K}_{sn} A_{sn})$$

Traditionally, scientists studying arousal often disagreed on a definition of arousal. Some viewed arousal as one large behavioral phenomenon, where others viewed it as a collection of multiple specific components. The concept outlined in the Arousal Equation unifies, for the first time, these two schools of thought and provides a precise, experimentally feasible manner of determining the arousal state of an animal.

Operational Definition and Quantitative Measures of Arousal

Generalized arousal can be experimentally determined and measured. By definition, an animal is considered to be more aroused if it has increased responsiveness to stimuli (S) from the environment (any sensory modality), increased voluntary motor (M) activity, and also increased ability to react emotionally (E). These three components encompass large classes of varied sensory modalities, which invariably give us a better measure of generalized CNS arousal. This is in sharp contrast to classic neuroscience, which sought to characterize the specific responses of an organism to a specific stimulus. Behavioral arousal is produced by select brain pathways and key populations of neurons being activated, eventually leading to changes in gene-expression patterns.

Using a statistical test called principal components analysis has enabled us to screen large amounts of data (e.g., behavioral responses to the three modalities of stimuli – sensory, motor and emotional reactivity – for each mouse) and uncover trends in the data. The use of this mathematical tool has revealed that generalized arousal accounts for approximately one-third (from 29.7 to 45%, depending on the experimental conditions) of all the data in arousal-related activities.

Application of Information Theory to Understanding the Neurobiology of Arousal

Introduced by Claude Shannon in the late 1940s, information theory, also known as Shannon's communication theory, establishes that the probability that an event will occur is inversely proportional to the information carried in that event. Consider, for example, the act of flipping a coin. There is a 50% chance of the coin landing heads up versus tails up; therefore, the possible information carried in this system is maximal – the more uncertain an outcome is, the more information transmittal there is. In contrast, if an event has a very high probability of occurring, there is very little room for doubt and very little information carried by the system. In neurobiological terms, the more unexpected the stimulus, the more dramatic the response (e.g., startle) will be. Likewise, the more nonintellectually stimulating environment, the less alert the animal will be and the less likely it is to respond to a stimulus (i.e., the less aroused it is).

Within the CNS, neuronal firing patterns are also governed by the principles of information theory. At any given time, there is a certain probability that a neuron will fire. The closer this probability is to 0.5 (or 50%), the more aroused the animal will be and the larger the repertoire of possible responses it can mount.

Neuroanatomical and Neurochemical Basis of Arousal

Proper arousal responses are critically important for nearly every aspect of our lives. Given this primal role, it is not surprising that nature has evolved ways to ensure that its regulation is safeguarded. This is accomplished by having multiple, redundant, and independent but cooperative mechanisms contributing to the overall arousal state.

Primordial Arousal-Promoting Neurons in the Arousal Crescent

A population of neurons in the reticular formation plays a key role in the regulation of arousal. These neurons originate in the ventral area of the brain stem, at the junction with the spinal cord, and extend bilaterally in a crescent-shaped form through the brain stem and eventually project ventrally into the thalamus and hypothalamus. This neuronal group is among the most evolutionarily conserved networks in animals, adding support to the hypothesis that it plays a crucial role in modulating arousal. These

neurons have a keen ability to sense and respond to a variety of arousing stimuli arising from within the body and the environment, including changes in blood pressure, pain, stress, and hormone levels. These neurons, mostly glutamatergic in nature, have ascending projections to the thalamus, hypothalamus, and basal forebrain, as well as descending projections to the spinal cord. These projections terminate at biologically relevant cell groups that are also involved in arousal control (e.g., the basal forebrain). Projections from this arousal crescent can travel via two routes: a ventral one and a dorsal one. The more ventral one is evolutionarily more primitive and terminates in the cholinergic basal forebrain. The more dorsal route terminates in the thalamus and hypothalamus and is more recent, evolutionarily speaking; it is especially well developed in primates.

Ascending Arousal Activating Pathways

In the 1940s and 1950s, Moruzzi, Magoun, and other neurophysiologists identified a key population of arousal-promoting neurons residing in the brain-stem reticular formation. These cholinergic neurons, found predominantly in the laterodorsal tegmental nucleus and the pedunculopontine tegmental nucleus, have extensive excitatory projections to the thalamus, which in turn project to widespread areas of the cerebral cortex. Acetylcholine release in both the thalamus and the cortex is highest during behaviorally aroused states of wakefulness and rapid eye movement (REM) sleep. In addition to the projections to the thalamus and eventually the cortex, brain-stem reticular neurons also project to the basal forebrain and hypothalamus, which contributes to stabilizing the aroused state.

Monoaminergic nuclei include the serotonergic dorsal raphe nuclei, adrenergic locus coeruleus, and dopaminergic substantia nigra in the brain stem, as well as the histaminergic tuberomammillary nucleus in the posterior hypothalamus. These nuclei all exhibit the highest discharge rates with maximal neurotransmitter release during wakefulness, lower discharge rates in non-REM sleep, and cease firing during REM sleep. The silencing of these neurons, and thus the absence of monoamine neurotransmission, plays a permissive role on REM sleep. These monoaminergic nuclei have ascending projections to the thalamus, hypothalamus, and cerebral cortex that favor states of behavioral arousal (e.g., wakefulness and REM sleep).

Descending Arousal Pathways

In antiphase to these arousal-promoting mechanisms, there are select brain regions that actively promote cortical synchronization, characteristic of non-REM sleep. One such brain nucleus is the hypothalamic ventrolateral preoptic nucleus (VLPO), whose neurons send inhibitory γ-aminobutyric acid (GABA) ergic and galaninergic projections to the monoaminergic cell groups, thus silencing them.

The lateral hypothalamus also contains neurons that are important for vigilance control. Wake-active orexin-producing neurons have been shown to have excitatory projections to monoaminergic cell groups. Significantly, monoaminergic projections to the VLPO express orexin receptors, and the lack of orexin binding at these receptors causes behavioral state instability associated with narcolepsy. In addition, intermingled with these neurons are melanin-concentrating hormone–GABA-containing neurons that are most active during REM sleep.

Neurochemicals Influencing Arousal

Acetylcholine Originally called vagusstoff (in reference to its release after stimulation of the vagus nerve), acetylcholine was the first neurotransmitter to be discovered. It is produced by the enzyme choline acetyltransferase and can bind to two classes of receptors: nicotinic and muscarinic. The major sources of acetylcholine in the CNS are the basal forebrain, pontine reticular formation, and medial septum. Cholinergic neurons from the reticular formation provide a major excitatory input, stimulating cortical activation, and the discharge patterns of these neurons are highest during behavioral aroused states.

Noradrenaline/norepinephrine In the CNS, norepinephrine is primarily produced by the locus ceruleus neurons, and these neurons promote behavioral arousal, primarily via their projections to the forebrain, brain stem, and spinal cord. Like other monoaminergic cell groups in the brain, locus ceruleus neurons discharge maximally during active waking, decrease their firing rates during non-REM sleep, and are silent during REM sleep. Norepinephrine inhibits VLPO sleep-promoting neurons by binding at α_2-adrenoreceptors. Also, significantly, norepinephrine increases the female sexual behavior lordosis by modulating the firing rates of ventromedial hypothalamic neurons, and sexual interactions between mice lead to increased norepinephrine release in the hypothalamus.

Serotonin Produced in the dorsal raphe nuclei of the brain stem, serotonin is a monoaminergic neurotransmitter that, like norepinephrine, promotes arousal. Serotonergic neurons fire maximally during waking (quiet waking), decrease their firing rate during non-REM sleep, and cease firing during REM sleep.

Many mood disorders, anxiety, and depression have been linked to deficiencies in serotonergic neurotransmission. Due to the large number of serotonin receptors (14) and their varied functions, studying the pharmacology of serotonin can be quite challenging.

Histamine Produced by tuberomammillary neurons in the posterior hypothalamus, histamine is a biogenic amine involved in vigilance control. With discharge patterns similar to serotonergic and noradrenergic neurons, maximal histamine release occurs during the waking state, and the administration of antihistamines is often accompanied by drowsiness. Like norepinephrine, histamine also increases sexual behavior in female mice by increasing the firing rates of ventromedial hypothalamic neurons. In addition, sleep-promoting VLPO region neurons send inhibitory GABAergic and galaninergic projections to histaminergic neurons, thus establishing a mutually exclusive, reciprocal interaction between the arousal and sleep-promoting brain regions.

Dopamine Neurons containing dopamine in the ventral periaqueductal gray matter have recently been identified as a major dopaminergic input to modulate arousal. These neurons have the highest discharge rates during wakefulness and have extensive projections to brain regions involved in sleep regulation. Dopaminergic neurons from the ventromedial nucleus and the substantia nigra may also contribute to promote arousal.

Orexin/hypocretin The recently discovered neuropeptides orexin/hypocretin are produced by a small population of neurons in the lateral hypothalamus perifornical area. The two peptides have the same precursor protein and, in addition to playing a role in stimulating appetite, they promote wakefulness. Orexinergic neurons have direct excitatory projections to the tuberomammillary nucleus, and a deficiency in either the orexin peptides or the receptors leads to behavioral state instability and narcolepsy.

Prostaglandin D_2 Prostaglandin D synthase catalyzes the isomerization of the common prostanoid precursor prostaglandin H_2 to the endogenous somnogen prostaglandin D_2. The involvement of prostaglandin D_2 in sleep regulation has been studied for the past 25 years. In fact, it was among the first endogenous sleep-promoting substances to be identified. Prostaglandin D_2 is thought to be involved in homeostatic sleep regulatory mechanisms because animals that lack prostaglandin D synthase or prostaglandin D_2 receptor fail to exhibit a sleep rebound in response to sleep deprivation. The somnogenic effects of prostaglandin D_2 are hypothesized to be mediated by the nucleoside adenosine, and infusion of prostaglandin D_2 into the subarachnoid space increases extracellular levels of adenosine.

Adenosine The nucleoside end product of energy metabolism, adenosine, has been shown to accumulate in a wake-dependent manner in the basal forebrain. The sleep-promoting effects of prostaglandin D_2 have been shown to be mediated by adenosine 2A receptors in the VLPO. Adenosine has been shown to disinhibit sleep-active VLPO neurons by blocking inhibitory inputs to these neurons and also by directly exciting these neurons. In addition, adenosine can inhibit wake-active neurons in the tuberomammillary nucleus. Interestingly, the most widely consumed arousal-promoting drug by humans, caffeine, works by blocking adenosine 2A receptors.

Steroid hormones Estrogens are produced by the ovaries and are secreted shortly before ovulation. The major role of estrogen is to prepare the female body for courtship, mating, and, eventually, maternal care. As such, estrogens modulate most of the neurochemicals and pathways mediating arousal to ensure that reproduction occurs. Examples are seen in cholinergic neurotransmission, in which estrogens not only increase the rate of acetylcholine synthesis but also increase the number of cholinergic receptors on target neurons. Similarly, estrogens increase the expression of genes involved in norepinephrine production and enhance the responses of dopamine and histamine. Conversely, the levels of the enzyme that produces the endogenous somnogen prostaglandin D_2 is decreased after estrogen treatment. In addition to the well-established roles of estrogens on sexual behaviors, there are many nonreproductive effects of estrogens on behavior, including anxiety, food intake, learning, memory, mood disorders, and sleep.

Integration of Signals

The neuroanatomical pathways discussed so far refer to mechanisms mediating arousal and sleep–wakefulness induction; there are also underlying circadian and homeostatic forces that determine when sleep can occur. The homeostatic drive for sleep, also known as process S, can be envisioned as an increased pressure that progressively gets larger the longer the animal is kept awake and that eventually reaches threshold levels, which trigger sleep. This mechanism is involved in modulating the day-to-day sleep distribution patterns, as well as sleep rebound in response to extended wakefulness. Homeostatic

sleep-regulatory molecules (e.g., prostaglandin D2, adenosine, and nitric oxide) accumulate in a wake-dependent manner and eventually stimulate sleep. Also communicating with these effector pathways is the circadian component, which determines when, along the 24-h light–dark cycle, sleep and arousal are most appropriate. In fact, the suprachiasmatic nucleus, the circadian pacemaker, has many direct and indirect projections to both sleep-promoting and arousal-promoting brain regions and thus contributes to the fine-tuning of arousal.

Amplified States of Arousal

Sex and Fear

Given the elaborate symphony of mechanisms and chemicals underlying arousal regulation, we quickly come to the conclusion that critical biological functions must rely on it. One such function is reproduction. The propagation of the species, our ultimate goal, is intimately dependent on proper levels of arousal at the appropriate times. The primal function that estrogens have is to prepare the body for reproduction. As such, redundant mechanisms have evolved that ensure that, once the hormonal milieu is optimal and ovulation is near, females exhibit the behaviors that attract mates and result in fruitful copulation. Although conceptually ideal, this scenario does not always occur. Imagine, if you will, that the animals are exposed to a stressful stimulus or are in an unfamiliar place; it may not be feasible to invest time and energy in reproduction when they should be worried, really, about survival. This is, in fact, what the literature tells us occurs in rodents. In the protective shelter of a running-wheel cage or a home cage in our arousal-monitoring equipment, after estrogen treatment the animals exhibited increased arousal (e.g., increased running-wheel activity). However, when the animals were tested in an open field or placed in an unfamiliar cage, the arousal behavior was quickly replaced by fear, resulting in lowered responsivity by the animal.

A fine line exists between which estrogen-mediated mechanisms are activated, depending on the environmental circumstances the animal finds itself; therefore, caution must be exercised when analyzing and interpreting such behavioral responses.

Significance of Studying Arousal

Arousal-related disorders affect people of all ages and genders. They comprise a wide spectrum of ailments, ranging from problems of vigilance control (e.g., narcolepsy) to hyperactivity disorders (e.g., attention-deficit hyperactivity disorder) and, in extreme cases, to devastating vegetative states, all of which can have detrimental and costly consequences.

Disorders of Reduced Arousal

Sleep disorders Forty million Americans are chronically affected by sleep disorders, and an additional 20–30 million report episodic sleep problems. Disorders of sleep range from the most common insomnias to sudden infant death syndrome to hypersomnias. Lack of research and adequate public education has led to a rising epidemic of sleep disorders, especially those associated with obesity (e.g., sleep apnea) and also to a chronically sleep-deprived society. Disorders of arousal mechanisms can be disastrous and costly, causing poor job performance, lack of attention in classrooms, and traffic accidents, all resulting from loss of vigilance.

Extreme hypo-arousal states The most extreme and devastating states in which arousal mechanisms are compromised can be seen in patients in vegetative, comatose, and stupurous states. Although the redundancy of neuroanatomical and neurochemical pathways minimize the possibility that this will occur, situations still exist in which the ability to be aroused is compromised. Only by understanding and characterizing these mechanisms can there be hope that someday it may be possible to stimulate certain key elements of this pathway (perhaps neurons in the arousal crescent) to reverse or ameliorate this condition.

Disorders of hyperarousal Cases of excessive arousal can be seen in patients suffering from attention-deficit hyperactivity disorder and posttraumatic stress disorder. Although they do not appear to be as debilitating as disorders of decreased arousal, being too aroused can interfere with a person's ability to function in society, for example, being able to learn in school or to focus at work. Likewise, the inability to repress unwanted detrimental memories in posttraumatic stress disorder can interfere with the healing process and the ability to resume normal activities.

Promising Future Lies Ahead in Arousal Research

Arousal mechanisms underlie all behavioral aspects of our existence. Although nature has done an astonishing job of ensuring that arousal mechanisms function optimally, as can be seen by the many independent but redundant pathways regulating arousal, situations still exist in which this system

is compromised. Only by knowing the intricacies of this regulation can researchers develop pharmacological and surgical therapies to reversing these problems.

See Also the Following Articles

Anxiety; Circadian Rhythms, Genetics of; Coping and Stress: A Lens and Filter Model; Drosophila Genes and Anoxia; Genetic Factors and Stress; Genetic Predispositions to Stressful Conditions; Heat Shock Genes, Human; Heat Shock Proteins: HSP60 Family Genes; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress; Expression Profiling of Stress Responsive Gene Patterns; Gene Environment Interactions in Early Development; Genetic Variation of HPA Axis Activity and Function in Farm Animals; Genetic Polymorphisms in Stress Response; Genetic Testing and Stress; Neuroticism Genetic Mapping of; Neuroticism Response to Stress, Genetic Mapping of Mice; Serotonin Transporter Genetic Modifications.

Further Reading

Borbely, A. A. (2001). From slow waves to sleep homeostasis: new perspectives. *Archives of Italian Biology* **139** (1–2), 53–61.

Garey, J., Goodwillie, A., Frohlich, J., et al. (2003). Genetic contributions to generalized arousal of brain and behavior. *Proceedings of the National Academy of Sciences USA* **100**(19), 11019–11022.

McEwen, B. S. (2004). *The end of stress as we know it*. Washington, DC: Dana Press/Joseph Henry Press.

Mong, J. A. and Pfaff, D. W. (2004). Hormonal symphony: steroid orchestration of gene modules for sociosexual behaviors. *Molecular Psychiatry* **9**(6), 550–556.

Pfaff, D. W. (2006). *Brain arousal and information theory: neural and genetic mechanisms*. Cambridge, MA: Harvard University Press.

Stress and Inflammation *See:* Inflammation.

Stress Effect of Assisted Reproduction

F M Helmerhorst, R Sibug and E R de Kloet
Leiden University Medical Center, Leiden, Netherlands

© 2007 Elsevier Inc. All rights reserved.

The Selye Alarm Reaction
Selye and Reproduction
Stress and Implantation

Glossary

Amenorrhoea	Absence of a menstrual period.
Anorexia nervosa	Relentless pursuit of thinness.
Anovulation	Absence of ovulation.
Blastocyst	Development of embryo with an inner (embryoblast) and outer (later placenta) cell mass.
COS	Controlled ovarian stimulation, by e.g. FSH in order to develop more than one ovarian follicle.
CRH	Corticotropin-releasing hormone.
Endometrium	Inner mucous membrane of the uterus.
flt-1	VEGF receptor.
GnRH	Gonadotropin-releasing hormone.
Gonadotropins	Tropic hormones (e.g. LH and FSH) that stimulate gonadal growth, steroidogenesis and gametogenesis.
Hypercorticolism	Increased concentration of cortisol.
KDR/flk-1	VEGF receptor.
LH	Luteinizing hormone.
FSH	Follicle-stimulating hormone.
hCG	Human chorionic gonadatropin.
hFSH	Human follicle-stimulating hormone.
hLH	Human luteinizing hormone.
HPA	Hypothalamic-pituitary-adrenal.
Perinatal	Around birth
VEGF	Vascular endothelial growth factor.

The Selye Alarm Reaction

In Knobil and Neill's *Physiology of Reproduction*, Ferrin defines stress "as the state in which the general adaptation syndrome has been evoked and in which

an increased demand requires the body to perform adaptive functions to re-establish normalcy" (2006: 2627). The three phases differentiated by Selye are the alarm reaction, the first phase in which the stressor affects the balanced system and then attempts to neutralize the stressor by catecholamine release; the second phase, in which the hypothalamic-pituitary-adrenal (HPA) axis acts to stabilize the situation by corticosteroid release; and, when this fails, the third and fatal stage.

Selye and Reproduction

The classic example of the Selye effect in reproduction is still observed when women develop amenorrhea on exposure to extremely violent and life-threatening situations, such as occurred during World War II. Corticotropin releasing hormone (CRH) inhibits gonadotropin-releasing hormone (GnRH) secretion in the hypothalamus, which leads to absence of follicle-stimulating hormone (FSH) and luteinizing hormone (LH) production and then to ovarian quiescence. Simultaneously, hypercortisolism is established. A similar mechanism of an activated HPA axis may lead to amenorrhea among endurance athletes and young women suffering from anorexia nervosa. However, in a consultation room of an infertility clinic, the perception of stress among normally cycling subfertile women is quite different from what we just have described as the Selye effect. The frequently expressed position of subfertile couples is that stress is involved in their infertility. The proof of their conviction as expressed by their friends and described in popular magazines is in the frequently occurring observation that there is a causal relation between taking a long weekend abroad and the desirable pregnancy. "Doctor, how can we avoid stress before and during the *in vitro* fertilization treatment?' asked a couple who was infertile for 5 years.

Three randomized controlled intervention trials illustrate the present situation between stress and assisted reproduction. One study by Kemeter and Feichtinger is low powered and flawed by methodology (e.g. no notification of the allocation concealment, doubts about randomization method), Tarabusi et al. and de Klerk et al. used surrogate endpoints (measuring depression and/or anxiety). From these sparse data we may at least conclude that a relationship between stress as defined in those publications and a hard clinical outcome after assisted procreation has not been properly investigated.

Stress and Implantation

In De Kloet et al., evidence is presented that suggests that stress-related events that occur between conception and the postnatal period can impact the offspring and permanently change their brains and behavior. The difficulties we encounter studying these processes and extrapolating from animal data to the human situation is illustrated by the fact that challenges that occur as early as ovarian hyperstimulation applied to facilitate blastocyst implantation may affect postnatal life.

Singleton pregnancies from assisted reproduction have a significantly worse perinatal outcome than nonassisted singleton pregnancies, although the prevalence of very low birth weight and very preterm birth is low. Very preterm birth is associated with later neuromotor and cognitive impairment, reduced school performance, and psychiatric morbidity. Young adults who are born very prematurely have different personality types than their term-born peers. This may be associated with an increased risk of psychiatric difficulties.

Subfertility cannot explain the risk of poor perinatal outcome after ovarian stimulation. Assisted reproduction is nearly always preceded by ovarian stimulation, in contrast to cryopreserved and thawed embryo transfer (ET). Because birth weight after cryopreserved and thawed ET is normal or even increased compared to control values, we might hypothesize that ovarian stimulation (with or without an IVF or intracytoplasmic sperm insemination procedure) is associated with low birth weight and might be considered a stress factor in the implantation process.

Blastocyst implantation is intimately associated with vascular permeability and angiogenesis. The vascular endothelial growth factor (VEGF) is a potent stimulator of these two processes. It has four isoforms in rodents, VEGF115, VEGF120, VEGF164, and VEGF188, which bind to the two tyrosine kinase receptors, flt-1 and KDR/flk-1. Our studies showed that VEGF120 and VEGF164 are highly expressed in the uterus. VEGF120, like VEGF164, showed a very distinct transcript signal and also exhibited a spatio-temporal distribution; the mRNA expression of VEGF, flt-1, and flk-1 correlates spatially and temporally with changes in angiogenesis and vascular permeability at implantation sites.

In our study, we tested the hypothesis that treatment with gonadotropins to induce controlled ovarian stimulation (COS) evokes a stress response, whose effects persist after conception. Adult female CD1 mice were treated with urinary human FSH (hFSH) and urinary human chorionic gonadotropin (hCG), recombinant hFSH and recombinant hLH, or with saline alone. The end point tested was the expression of VEGF120 during the mouse peri-implantation period. Urinary gonadotropic treatment caused lower levels of VEGF120, flt-1, and flk-1 mRNA levels, reduced the size of the embryo implantation site,

delayed implantation and prolonged the gestational period; both urinary hFSH and urinary hCG contributed to the adverse effects. Recombinant gonadotropin treatment did not alter any of the parameters studied. In addition, our recent gene-expression profiling study using microarray technology showed that urinary gonadotropins, but not recombinant gonadotropins, exerted differential effects on gene-expression patterns during the murine peri-implantation period.

The CRH system is another factor that is hypothesized to play a role systemically and/or locally in blastocyst implantation. Studies in rodents have shown that CRH expression is 3.5-fold higher in the implantation sites than interimplantation sites and that the injection of a CRH receptor 1 (CRH-R1) antagonist resulted in a 50–70% reduction in the number of implantation sites. Moreover, the administration of antibody against CRH inhibits the attachment of the blastocyst to the implantation site. The activation of CRH receptors during the implantation phase has been demonstrated in the human endometrium. We, therefore, tested in the same system whether the urinary and recombinant gonadotropins influence the CRH system during the mouse peri-implantation period. The levels of CRH and CRH-R1 expression were not influenced, suggesting that the CRH system is not involved in mediating the adverse effects of urinary gonadotropins on the implantation process.

The urinary gonadotropins did not show a significant effect on birth weight, but we observed a positive correlation between birth weight and length of gestation, suggesting that the longer period of gestation allowed time for the fetuses to compensate for the delayed growth and achieve comparable birth weights.

Further Reading

Brundu, B., Loucks, T. L., Adler, L. J., et al. (2006). Increased cortisol in the cerebrospinal fluid of women with functional hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* **91**, 1561–1565.

De Klerk, C., Hunfeld, J. A. M., Duivenvoorden, H. J., et al. (2005). Effectiveness of a psychosocial counselling intervention for first-time IVF couples: a randomized controlled trial. *Human Reproduction* **20**, 1333–1338.

De Kloet, E. R., Sibug, R. M., Helmerhorst, F. M., et al. (2005). Stress, genes and the mechanism of programming the brain for later life. *Neuroscience and Biobehavioral Reviews* **29**, 271–281.

Ferin, M. (2005). Stress and the reproductive system. In: Knobil, E. & Neill, J. D. (eds.) *Physiology of reproduction* (3rd edn., pp. 2627–2694). Boston: Elsevier.

Helmerhorst, F. M., Perquin, D. A. M., Donker, D., et al. (2004). Perinatal outcome of singletons and twins after assisted conception: a systematic review of controlled studies. *British Medical Journal* **328**, 261–265.

Kapiteijn, K., de Bruijn, C. S., de Boer, E., et al. (2006). Does subfertility explain the risk of poor perinatal outcome after ivf and ovarian hyperstimulation? Manuscript.

Kemeter, P. and Feichtinger, W. (1986). Prednisolone supplementation to Clomid and/or gonadotropin stimulation for in-vitro fertilization: a prospective randomized trial. *Human Reproduction* **7**, 441–444.

Sibug, R. M., Datson, N., Tijsssen, A. M. I., et al. (2006). Effects of urinary and recombinant gonadotropins on gene expression profiles during the murine peri-implantation period. *Journal of the Society for Gynecologic Investigation* **13**(2, supplement), 357A.

Sibug, R. M., de Koning, J., Tijssen, A. M., et al. (2005). Urinary gonadotropins but not recombinant gonadotropins reduce expression of VEGF120 and its receptors flt-1 and flk-1 in the mouse uterus during the peri-implantation period. *Human Reproduction* **20**, 649–656.

Sibug, R. M., Helmerhorst, F. M., Tijssen, A. M. I., et al. (2002). Estrogen reduces vascular endothelial growth factor 164 expression in the mouse nucleus paraventricularis of the hypothalamus. *Neuroscience Letters* **333**, 199–202.

Sibug, R. M., Helmerhorst, F. M., Tijssen, A. M. I., et al. (2002). Gonadotropin stimulation reduces VEGF (120) expression in the mouse uterus during the peri-implantation period. *Human Reproduction* **17**, 1643–1648.

Tarabusi, M., Volpe, A. and Facchinetti, F. (2004). Psychological group support attenuates distress of waiting in couples scheduled for assisted reproduction. *Journal of Psychosomatic Obstetrics and Gynecology* **25**, 273–279.

Stress Effects, Overview

A Steptoe
University College London, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A Steptoe, volume 3, pp 510–511, © 2000, Elsevier Inc.

Glossary

Coping responses	Cognitive and behavioral responses mobilized in an effort to manage stressful transactions.
Individual-response specificity	The proclivity of an individual to respond to a range of challenges with a consistent profile of physiological activation.

The effects of stress are manifest in four distinct domains: physiology, behavior, subjective experience, and cognitive function. The physiological effects of stress include alterations in neuroendocrine, autonomic nervous system, and immune function and are of great interest to biomedical researchers because of their implications for disease. The behavioral manifestations of stress include aggressive responses, freezing, and displacement activities in animals; health-related behaviors such as smoking and alcohol consumption in humans; and deficits in the performance of complex tasks that are particularly relevant to public safety in areas such as transport and pollution control. Subjective experiences during stress include distress and feelings of dissatisfaction and anger, together with anxiety and depressive responses, which may contribute to psychiatric ill health. The cognitive manifestations of stress include alterations in information processing and attentional and memory functions, which can, in turn, influence performance and decision-making capability.

Stress effects are patterned both within and across these four domains. Attempts to single out a gold-standard measure that provides a consistent indicator of stress responses in all situations have proved unsuccessful. For example, research in the tradition initiated by Hans Selye has emphasized the hypothalamic-pituitary-adrenocortical axis and the output of corticosteroids as central elements of physiological stress responses. Although it is true that this pathway is very sensitive to stressors, it does not necessarily relate closely to other responses. Thus, differences in the stress-induced inhibition of immune responses with varying stressor predictability and controllability have been found to be independent of corticosteroid responses. Particularly low levels of cortisol have been recorded in victims of posttraumatic stress disorder, suggesting that some manifestations of stress are independent of heightened adrenocortical function. Similarly, it is commonplace in human studies to find weak or negligible correlations between the magnitude of subjective distress and the activation of physiological indices of stress.

Several factors appear to be important in understanding these variations in stress effects. The first is that the pattern of change depends on the nature of the stressful experience and the coping responses mobilized to manage the situation. Thus, naturalistic studies of antagonistic encounters in animals have drawn a distinction between active behavioral responses involving fighting and escape that are associated with sympathetic nervous system activation and more passive conservation–withdrawal behaviors that tend to be linked with parasympathetic responses. At the level of subjective disturbance, depressed mood tends to be elicited by loss events and experiences, whereas anxiety is more characteristic as a response to threatening occurrences.

Second, stress effects vary in their time course and duration. Coping responses that are effective acutely may become maladaptive when stimulation is prolonged. Some important differences in responses emerge only with chronic aversive stimulation, whereas other effects only arise after stimulation is terminated. For instance, studies of human information processing have shown that performance may be maintained during stress exposure due to the exertion of additional effort and that only later manifestations of disturbance emerge. Chronic or repeated stimulation may lead to the failure of regulatory physiological feedback systems, with damaging consequences. Thus, research on socioeconomic status and stress responsivity indicates that differences across the social gradient are more apparent in rates of post-stress recovery than in the magnitude of acute biological responses.

When the implications of stress research for health and well-being are considered, the variations in response between individuals are striking. Two people exposed to the same threatening situation may differ substantially in the magnitude and duration of stress responses, and stress-related health problems may emerge in several contrasting ways both physically and mentally. Some of these variations result

from differences in temperament, social resources, and the effectiveness of the coping responses that the individual brings to bear on the stressful transaction. Others relate to differences in vulnerability, which may arise through both genetic and experiential factors. Genetic variations both in central neurotransmitter metabolism and in peripheral receptor sensitivity contribute to variations in biological response. Animal studies have shown that experiences early in life, such as gentle handling or maternal separation, are associated with decreases or increases in neuroendocrine responses to stressful stimulation in adult life. Clinical investigations of adult psychiatric disorder have uncovered variations in childhood trauma and abuse that sensitize the individual to subsequent stressful events.

Another relevant factor is individual-response specificity. Identified through research on acute psychophysiological responses, this is the tendency for some people to react with a consistent physiological response profile across a range of challenges. One person may be particularly responsive in terms of blood pressure and will show exaggerated blood pressure responses to a range of situations, whereas another is more responsive hormonally or gastrointestinally. Such variations may signal the existence of a special vulnerability or lack of resistance in a particular biological system. Background factors such as nutritional status, physical fitness level, age, gender, and socioeconomic status all serve to moderate the extent to which stress responses may have long-term adverse effects. Hence, the prediction of stress effects in any particular case depends not only on assessing the responses in the four domains with appropriate measures but also on understanding the background factors that predispose to poor adaptation or alternatively bolster the individual's resilience.

Further Reading

Cannon-Bowers, J. A. and Salas, E. (eds.) (1998). *Making decisions under stress*. Washington, DC: American Psychological Association.

Charmandari, E., Tsigos, C. and Chrousos, G. (2005). Endocrinology of the stress response. *Annual Review of Physiology* **67**, 259–284.

Charney, D. S. (2004). Psychobiological mechanisms of resilience and vulnerability: implications for successful adaptation to extreme stress. *American Journal of Psychiatry* **161**, 195–216.

Marmot, M. G. and Wadsworth, M. E. J. (eds.) (1997). *Special issue on fetal and early childhood environment: long-term health implications British Medical Bulletin* **53**(1).

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* **338**, 171–179.

Pacak, K. and Palkovits, M. (2001). Stressor specificity of central neuroendocrine responses: implications for stress-related disorders. *Endocrine Reviews* **22**, 502–548.

Raison, C. L. and Miller, A. H. (2003). When not enough is too much: the role of insufficient glucocorticoid signaling in the pathophysiology of stress-related disorders. *American Journal of Psychiatry* **160**, 1554–1565.

Steptoe, A. (in press). Psychophysiological contributions to behavioral medicine and psychosomatics. In: Cacioppo, J. T., Tassinary, L. G. & Bernston, G. (eds.) *The handbook of psychophysiology* (3rd edn.). New York: Cambridge University Press.

Steptoe, A. and Marmot, M. (2005). Socioeconomic position and coronary heart disease: a psychobiological perspective. In: Waite, L. J. (ed.) *Aging, health, and public policy*, pp. 133–150. NewYork: Population Council.

Suomi, S. J. (2003). Gene-environment interactions and the neurobiology of social conflict. *Annals of the New York Academy of Sciences* **1008**, 132–139.

Yehuda, R. (2002). Current status of cortisol findings in post-traumatic stress disorder. *Psychiatric Clinics North America* **25**, 341–368.

Stress Generation

J E Roberts
University at Buffalo, The State University of New York, Buffalo, NY, USA
J A Ciesla
Vanderbilt University, Nashville, TN, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J E Roberts and J A Ciesla, volume 3, pp 512–517, © 2000, Elsevier Inc.

Life Stress in Psychopathology
Generation of Stressful Life Events
Generation of Stress Perceptions and Threat
Issues in Stress Assessment

Glossary

Appraisal	A cognitive process that involves evaluating the degree of threat associated with a life stressor, as well as the implications of that stressor.
Camberwell life events and difficulties schedule	An interview-based procedure for evaluating the presence, severity, and origins of stressful events and difficulties.
Diatheses	Vulnerability factors that increase the stressfulness and negative consequences of life events.
Difficulties	Ongoing stressful situations that persist for at least 3 months.
Independence	The degree to which a person's behavior or characteristics might have brought about a stressful life event.
Matching events	Acute major negative events that arise from or are thematically related to ongoing chronic stressors or role conflicts.

Life Stress in Psychopathology

Links between life stress and mental disorders such as depression have been drawn throughout history. For example, in the seventeenth century Robert Burton described how melancholia often was precipitated by environmental adversities, including interpersonal losses. Indeed, the very phenomenology of depression involves feelings of loss and ruminations on the hopelessness of one's life situation. When depressed, one experiences the world as impoverished and as imposing impossible demands. In such a frame of mind, ordinary hassles become insurmountable obstacles, deprivations, or failures over which people despair. Thus, depression itself magnifies the apparent stress and strain of life.

Investigators have typically examined the role that life stress might play in the onset and maintenance of mental disorders. In other words, the focus has been on how life stress can result in psychopathology. However, it is also clear that life stress often results from individuals' behavior or characteristics. Many stressful life events and difficulties, therefore, are not random occurrences that happen to befall the unfortunate, but rather are at least in part created by the same individuals who suffer from them.

Likewise, certain individuals tend to react to negative life events with greater feelings of stress and distress than others. Increasingly, investigators are exploring the manner in which characteristics of the person and the environment can breed stressful life events and difficulties, as well as the more subjective stress reactions to those events. This generated stress, in turn, potentially contributes to other adverse psychological outcomes, such as episodes of depression. Although this is largely studied in the context of depression, an important area of future research involves determining whether other forms of psychopathology also contribute to stress generation.

As suggested earlier, stress generation can involve two separate processes. First, characteristics or behavior of the individual can contribute to the occurrence of actual life events. In other words, for various reasons, some people are more prone to encountering negative life events and difficulties than others. Second, characteristics of the individual can contribute to the level of threat or perceived stress associated with particular life events. In other words, given seemingly identical life events, some people will experience more severe subjective stress reactions than others.

Generation of Stressful Life Events

It is clear that people tend to select or even forge environments that reflect their personalities and genetic predispositions. Individuals with various temperaments create social-environmental niches that are consistent with these characteristics, and in turn these environments are likely to be associated with various types of life events. For example, studies have demonstrated that persons with higher levels of the personality characteristic neuroticism are more prone to experiencing negative life events, whereas those who are higher in extroversion are more prone to experiencing positive life events. Likewise, it is

quite possible that cognitive diatheses associated with depression, such as negative attributional styles or dysfunctional beliefs about interpersonal relationships, can lead to future stressful life events. For example, an individual who is cognitively biased to interpret ambiguous social encounters as rejection is likely to create the very abandonment so feared. The person's negative interpretations are likely to lead to conflict and rejection. In a similar vein, there is strong evidence that individuals genetically predisposed to depression are at heightened risk for negative life events. For example, the nondepressed biological relatives of depressed persons have been shown to be at greater risk of experiencing major life stressors than relatives of nondepressed individuals. Several quantitative behavioral genetic studies have confirmed this genetic contribution to stressful life events. Overall it is apparent that negative life events are not entirely randomly distributed and, instead, tend to cluster in certain persons. Furthermore, these individuals seem to be the same persons who are most vulnerable to psychological disorders, particularly depressive episodes, following life stress.

Brown and colleagues were among the first to explore the role that psychiatric symptomatology might play in generating stressful life events. Initial efforts by this research team were directed at establishing that negative life events could play a causal role in the development of subsequent depressive episodes. It was therefore critical that these investigators focused on events that clearly arose independently of psychiatric symptomatology. Across a number of studies, these and other researchers demonstrated that major acute stressful life events that were independent of depression itself (as opposed to events such as hospitalization for depression or losing one's job because of concentration difficulties) contributed to the onset of episodes of this disorder. However, these investigators also established that individuals with subclinical depressive conditions, as well as other psychiatric conditions, were at greater risk of encountering environmental adversities than those who were without symptomatology. In other words, individuals with subclinical depression or other psychiatric conditions were more likely to be faced with major acute stressful life events than were asymptomatic individuals. These acute stressors, in turn, were associated with the onset of episodes of clinical depression. Interestingly, subclinical depression and other psychiatric conditions failed to predict future depressive episodes in the absence of life stressors. Apparently, stress generation was largely responsible for the exacerbation of depressive symptoms, leading subclinical experiences to develop into full depressive episodes.

Furthermore, subclinical depression and other psychiatric conditions were associated with the generation of life events that are particularly depressogenic. Specifically, these conditions bred acute life events that matched either ongoing life difficulties (e.g., a conflictual marriage) or role conflicts (e.g., a conflict between one's career and family). These types of matching events have proven to be extremely potent contributors to the onset of episodes of clinical depression. In other words, matching events seem to be considerably more likely to result in clinical depression than other types of negative life events. Unfortunately, depression-prone individuals appear to generate the very types of stressors that are most likely to exacerbate their psychiatric conditions.

In a highly influential paper, Hammen also suggested that stress generation might be an important mechanism through which depression can lead to future depression. Across a number of studies based on adult, adolescent, and child samples, Hammen and colleagues have demonstrated that depression-prone individuals experience a greater frequency of subsequent life stressors compared to those with other psychiatric conditions, those with medical illnesses, and healthy control participants. Interestingly, this research suggests that interpersonal stressors, in particular, are generated by depression-prone individuals. Interpersonal stressors largely include conflicts and tension between people. Importantly, Hammen's work has demonstrated that stress generation can take place even during periods of remission, suggesting that factors beyond depressive symptomatology itself are likely to be at work. In other words, stress generation appears to be driven by some correlate or correlates of depression rather than depressive symptomatology itself.

What might these correlates be? Studies have explored aspects of personality pathology and interpersonal functioning, such as poor interpersonal problem solving and excessive reassurance seeking. For example, symptoms of cluster A (odd and eccentric) and cluster B (dramatic and erratic) personality pathology have been found to contribute to the generation of interpersonal stress, which in turn predicted the development of depressive symptomatology. Other research has focused on two personality styles implicated in vulnerability to depression, sometimes referred to as interpersonal relatedness (a style marked by excessive dependence on relationships) and self-definition (a style marked by excessive reliance on achievement and individuality in maintaining self-esteem), and has found that self-definition is particularly related to stress generation. In terms of interpersonal functioning, one study found that the generation of minor social stressors mediated the association between excessive

reassurance seeking and the development of depressive symptoms, whereas another found that the generation of interpersonal stressors mediated the association between poor interpersonal problem solving and subsequent depressive symptoms. Another line of research conducted by Kendler and colleagues has found that neuroticism and genetic liability to depression contribute to stress generation, whereas childhood sexual abuse and variation in the serotonin transporter gene increase individuals' sensitivity to the depressogenic effects of life stressors. Together, these studies suggest that personality pathology, interpersonal functioning, and some forms of genetic liability are associated with future depression by increasing the likelihood that individuals encounter stressful life events (particularly interpersonal stressors), whereas traumatic stress early in life and other types of genetic variability increase stress sensitivity. It will be important for future work to determine whether these characteristics contribute to stress generation and sensitivity on their own, or whether their effects become more potent in combination with pre-existing depressive symptomatology.

In addition to these characteristics of the person, it is likely that characteristics of the person's social environment are also related to stress generation. Brown and colleagues have documented how other ongoing aspects of the social world are associated with risk for future acute life events. These investigators developed the concept of life structure as being important in understanding the origins of life stress. In many cases, individuals who are faced with various forms of deprivation during childhood (e.g., parental loss, abuse, or neglect) become locked into an ongoing conveyor belt of environmental adversity and psychological vulnerability. These childhood environmental experiences put individuals at risk for further negative environmental experiences during later stages of life. For example, parental lack of care, neglect, and abuse during childhood can lead to teen pregnancy and marriage to an undesirable partner. In turn, these experiences set the stage for a variety of additional acute stressful life events and difficulties that are likely to occur in the future, such as marital violence, economic hardship, and infidelity. Simultaneously, this conveyor belt of adversity would continuously foster and reinforce psychological vulnerabilities, such as feelings of helplessness and low self-esteem, making it increasingly difficult to cope with stressors.

In summary, acute negative life events are more likely to be encountered by individuals with certain characteristics. In particular, this type of stress generation is associated with some forms of genetic predisposition, neuroticism, depression, and, quite possibly, cognitive characteristics that are thought to increase risk for depression. Further, these generated life events appear to be characterized by interpersonal conflict and a resonance with other ongoing life stressors or role conflicts. It is presently unclear why depression-prone individuals are at risk for generating life stress. Important directions for future work include examining theoretically plausible mediators and moderators of this effect. For example, Hammen and colleagues have begun to explore the possible roles of interpersonal functioning, comorbidity, and personality pathology, while Brown and colleagues have focused on the individual's unfolding social-environmental context. Future work also needs to examine stress generation in psychological disorders beyond depression.

Generation of Stress Perceptions and Threat

Individuals differ in the degree to which they experience negative reactions to seemingly identical life events, i.e., stress reactions are determined not only by the nature of the event, but also by the nature of the individual. Lazarus and Folkman proposed that stress itself is best thought of as a process in which the individual actively mediates the relationship between life events and psychological distress. In response to a life event, individuals are thought to first engage in a primary appraisal of the potential negativity or threat. Different individuals are likely to appraise the same event in quite different ways. One person may view a job demotion as an irreversible career blemish, while another may view this event as only a minor setback. Subsequent to this initial threat appraisal, individuals are thought to engage in another assessment. This secondary appraisal involves further evaluation of the situation and the degree to which the person has the resources available to effectively deal with the event. In a number of ways, this involves cognitive, emotional, and social aspects of the individual. Finally, following primary and secondary appraisal, different coping strategies can be initiated. Individuals are known to differ in their coping efficacy and typical coping styles, both of which are likely to effect psychological adjustment to life stress. In describing this stress process, Lazarus has been influential in identifying multiple ways in which properties of the individual can affect vulnerability to life stress.

Lazarus' model highlights the importance of individual differences in appraisal processes as contributing to the generation of stress reactions, but what underlies and influences these appraisals? From a theoretical perspective, other types of vulnerability

(so-called diatheses) may contribute to these negative appraisals. Diathesis stress models posit that certain characteristics increase the stress that individuals experience in response to negative life events. One such vulnerability factor that has received empirical support is attributional style.

According to Abramson and colleagues, individuals are motivated to understand and make sense of the causes of life events and tend to have relatively stable characteristic attributional styles. Persons who attribute negative life events to causes that are global (i.e., have far-reaching implications), stable (i.e., are predictive of future negative events), and internal (i.e., are caused by the self) are more likely to experience psychological distress than individuals who attribute negative life events to causes that are specific (i.e., are limited in life domain), unstable (i.e., are not likely to occur again), and external (i.e., are caused by the environment). While this is true of negative events, the opposite pattern is considered dysfunctional in the context of positive life events.

In addition to negative attributional styles, rigid, inappropriate, and perfectionistic beliefs about the self and the world, known as dysfunctional attitudes, likely contribute to the generation of stress perceptions and threat. Many of these beliefs concern contingencies of self-worth that are unrealistic in nature (e.g., one must be perfect or admired by all in order to be a person of worth or deserving of love). These cognitions are thought to interact with life stressors to result in psychological disorders. In this sense, a negative life event may be experienced as distressing in itself, but may also be considered by the individual to be an indication of a lack of self-worth. Through these beliefs, an event such as being laid off due to a factory closing may become a sign of one's personal incompetence and lack of value, generating more negative appraisals and perceived stress than would be generated by someone without these dysfunctional beliefs. In this way, the manner in which an individual's self-esteem is maintained and structured influences how much perceived stress is generated by particular classes of negative life events.

It is also apparent that environmental factors can contribute to these appraisal processes and, consequently, to the degree of stress generated by the occurrence of negative life events. For example, individuals with inadequate social support are likely to have more severe stress reactions than those with more supportive close relationships. Brown and colleagues have found that individuals who were "let down" during the negative event itself were more likely to develop episodes of depression than those who received expected crisis support. In the former situation, individuals' positive expectations about their close relationships would be shattered, leading to disappointment and a reevaluation of the nature and quality of these relationships. Consequently, the event and its surrounding context would likely be appraised as considerably more threatening than if support had been forthcoming.

Likewise, individuals who are already faced with significant life stressors, particularly severe ongoing difficulties, are more likely to suffer from more severe stress reactions in response to additional negative life events than those without these additional burdens. Making matters more complicated, Coyne and Wiffen have argued that it is likely that these environmental factors are closely associated with the cognitive vulnerabilities discussed previously. For example, to some extent attributions about the implications of negative interpersonal events or dysfunctional beliefs about relationships likely arise from conflictual interpersonal relationships, such as unstable marriages. It becomes less clear whether it is primarily the cognitive characteristic (e.g., negative attributional style or dysfunctional beliefs about relationships) or the environmental factor (e.g., conflictual marriage) that contributes to the generation of stress reactions following an acute negative life event (e.g., a major argument with one's spouse).

In summary, characteristics of the individual and his or her social environment contribute to the degree of stress and distress that is generated in response to negative life events. These characteristics include appraisal processes that, in turn, are likely to be driven by underlying cognitive diatheses (such as negative "attribution styles" and dysfunctional beliefs), as well as environmental characteristics (such as inadequate social support and chronic stressors). An important area of future investigation involves attempting to disentangle the overlapping roles of these cognitive and environmental characteristics.

Issues in Stress Assessment

The notion of stress generation mandates that careful attention be paid to a number of issues in the assessment of life stress. First, it becomes crucial to examine the dimension of event independence. Presumably, generated events are those that were at least somewhat dependent on the individual's behavior or characteristics. Second, as previously discussed, stress generation involves two different processes: the first reflects the generation of actual negative life events, whereas the second reflects the generation of stress perceptions and reactions. Assessment procedures must be able to distinguish between these two processes. Third, it is becoming clear that in terms of objective events, stress generation is specific to certain classes of events, namely, events that match ongoing problems or that involve interpersonal

conflict. Assessment procedures need to be able to cleanly distinguish these classes of events from others.

In contrast to the stress generation studies conducted by Brown, Hammen, and their colleagues, the vast majority of research on the role of life stress in depression has relied on self-report checklists to assess life events. Unfortunately, self-report measures of life stress likely combine both relevant and irrelevant aspects of the stress construct, making it difficult to advance knowledge in this area. First, self-report measures of life events are unable to partition the occurrence of objective life events from the perception of life stress. As a result, it is unclear whether a participant's score is the result of actual changes in the environment or the perception of changes. A second major limitation of self-report life event measures is that they are unable to assess the degree to which events were influenced by the person or by other independent causes. Clearly, in order to investigate the hypothesis that individuals contribute to their own stress, it is necessary to establish that a given stressor was the result of an individual's behavior and was not completely independent or random. Finally, self-report measures make it difficult, if not impossible, to classify events into particular domains (e.g., interpersonal and matching versus nonmatching) in order to examine specific types of events.

Importantly, Hammen and colleagues' work suggests that depression generates interpersonal conflict events, but not other types of events. As such, these interpersonal conflict events should be examined separately from other types of events. It is notoriously difficult to make these distinctions on the basis of self-report checklists. For example, events that seemingly reflect noninterpersonal experiences (such as loss of one's job) often can involve an interpersonal conflict (such as a fight with one's boss). More in-depth probing is required to determine whether or not an event had significant interpersonal overtones.

Clearly, advancements in this area of inquiry are going to require the use of more sophisticated interview-based procedures that can more cleanly assess particular dimensions and domains of stressful life events. The Camberwell Life Events and Difficulties Schedule (LEDS) is the most well-developed and methodologically sound system. The LEDS first involves careful questioning about potential events in a variety of different life domains, as well as in-depth probing about the context surrounding each event. Events are then rated on a number of dimensions, including threat and independence, by a team of individuals who did not conduct the interview and who are kept blind to the psychological reactions to the events. In order to evaluate the likely individual meaning and importance of the event, ratings of degree of threat (stress) are made on the basis of the context surrounding the event. So, for example, a pregnancy would be rated quite differently if it occurred in the context of a happy marriage and good financial resources than it would if it occurred in the context of marital violence, financial hardship, and poor living quarters. Of particular importance is the fact that the LEDS includes more than 5000 examples of different events that help anchor ratings of threat and independence. In addition, guidelines are provided to help determine what counts as an event. These guidelines set thresholds for inclusion and exclusion. Consequently, they help prevent individuals who report a number of very minor and trivial incidents from receiving higher scores than those who experienced the same incidents, but who failed to mention them during the interview. The LEDS is currently the only interview that includes these important features.

The notion of stress generation implies that the stressful life event in part resulted from the person's behavior or characteristics. As such, it crucial that the dimension of life event independence is examined. The LEDS classifies events on a 12-point scale that includes categories that vary in terms of their independence of the person's behavior. Point 1 refers to events that are completely independent, such as accidents occurring to family members; point 2 refers to events for which it is impossible to completely rule out some aspect of the subject's behavior in leading to the event, e.g., a close friend deciding to move to another city; point 3 refers to events that possibly arose through the subject's behavior, but is unlikely, e.g., a child's academic difficulties; point 4 refers to physical illnesses of the subject; point 5 refers to events in which the subject consented or complied with external circumstances, e.g., agreeing to care for a sick relative; point 6 refers to intentional acts of the subject, e.g., deciding to return to school; point 7 refers to events that probably arose as a result of the subject's negligence, e.g., an automobile accident while subject was intoxicated; point 8 refers to events involving arguments or breaking off contact with someone after tension; point 9 refers to breaking off contact with someone without a precipitating argument or period of tension; point 10 refers to the subject's love events; point 11 refers to the subject's partner's love events, e.g., a partner's affair; and point 12 refers to events that are clearly the result of depression, e.g., fatigue, irritability, and concentration problems contributing to reprimands from one's boss and, eventually, to being fired.

In summary, research is beginning to document the ways in which individuals prone to psychological disorders, particularly depression, contribute to the stress that they experience in their lives. In turn, this generated stress might effect the onset, maintenance,

and recurrence of these conditions. Our developing understanding of these complicated transactional pathways suggests that more refined evaluations of life stress are required to advance knowledge in this area. In-depth interview-based approaches, such as the LEDS, are required to evaluate life events in terms of their severity, independence, content, and relation to other ongoing life problems.

See Also the Following Articles

Life Events and Health; Life Events Scale; Genetic Predispositions to Stressful Conditions; Psychosocial Factors and Stress.

Further Reading

Abramson, L. Y., Alloy, L. B. and Metalsky, G. I. (1989). Hopelessness depression: a theory-based subtype of depression. *Psychological Review* **96**, 358–372.

Brown, G. W. and Harris, T. O. (1978). *The social origins of depression*. New York: Free Press.

Brown, G. W. and Harris, T. O. (eds.) (1989). *Life events and illness*. New York: Guilford Press.

Coyne, J. C. and Whiffen, V. E. (1995). Issues in personality as diathesis for depression: the case of sociotropy-dependency and autonomy-self-criticism. *Psychological Bulletin* **118**, 358–378.

Daley, S. E., Hammen, C., Davila, J. and Burge, D. (1998). Axis II symptomatology, depression, and life stress during the transition from adolescence to adulthood. *Journal of Consulting and Clinical Psychology* **66**, 595–603.

Davila, J., Hammen, C., Burge, D., Paley, B. and Daley, S. E. (1995). Poor interpersonal problem solving as a mechanism of stress generation in depression among adolescent women. *Journal of Abnormal Psychology* **104**, 592–600.

Hammen, C. (1991). Generation of stress in the course of unipolar depression. *Journal of Abnormal Psychology* **100**, 555–561.

Kendler, K. and Kartowski-Shuman, L. (1997). Stressful life events and genetic liability to major depression: genetic control of exposure to the environment? *Psychological Medicine* **27**, 539–547.

Lazarus, R. S. and Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Monroe, S. M. and Roberts, J. E. (1990). Conceptualizing and measuring life stress: problems, principles, procedures, progress. *Stress Medicine* **6**, 209–216.

Monroe, S. M. and Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. *Psychological Bulletin* **110**, 406–425.

Potthoff, J. G., Holahan, C. J. and Joiner, T. E. (1995). Reassurance seeking, stress generation, and depressive symptoms: an integrative model. *Journal of Personality and Social Psychology* **68**, 664–670.

Roberts, J. E. and Monroe, S. M. (1994). A multidimensional model of self-esteem in depression. *Clinical Psychology Review* **14**, 161–181.

Scarr, S. and McCartney, K. (1983). How people make their own environments: a theory of genotype – environment effects. *Child Development* **54**, 424–435.

Stress Hyporesponsive Period

S Levine
University of California, Davis, CA, USA
E R de Kloet
Leiden/Amsterdam Center for Drug Research, Leiden, Netherlands
G Dent
AstraZeneca Pharmaceutical Co., Wilmington, DE, USA
M S Schmidt
Max Planck Institute of Psychiatry, Munich, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Levine, G Dent, and E R de Kloet, volume 3, pp 518–526, © 2000, Elsevier Inc.

Stress Hyporesponsive Period

Stress Hyporesponsive Period, the Adrenal, and Corticosterone

Stress Hyporesponsive Period and the Pituitary

Stress Hyporesponsive Period and the Brain

Corticosteroid Feedback

Conclusion

Glossary

Maternal deprivation	Separation of mother and infant during the stress hyporesponsive period, which needs to last for at least 8 h for immediate and persistent effects on the neuroendocrine regulation of the hypothalamic-pituitary-adrenal axis.
Stress hyporesponsive period (SHRP)	A period of reduced adrenal corticosterone and pituitary adrenocorticotropic hormone release in response to stress, lasting in the rat from postnatal day 4 to 14 and in the mouse from postnatal day 1 to 12.

Based primarily on the pioneering work of the late Hans Selye, the stress response has become somewhat synonymous with the release of hormones from the pituitary and the adrenal glands. Thus, in most adult mammals stimuli presumed to be stressful result in a systematic release of adrenocorticotropic hormone (ACTH) and the subsequent secretion of glucocorticoids from the adrenal. This simplistic view of the pituitary-adrenal axis as first described by Selye has been elaborated on extensively. The regulation of the stress hormones clearly involves specific peptides synthesized and stored in the brain, corticotropin-releasing factor (CRF) and arginine vasopressin (AVP), and brain-derived neurotransmitters, such as norepinephrine. Thus, the brain must be included as a critical stress-responsive system. However, the sequences of responses observed consistently in the adult are in many ways very different in the developing organism. Abundant evidence indicates that the rules that govern the activity of the hypothalamic-pituitary-adrenal (HPA) axis in the adult are very different in the neonate.

Stress Hyporesponsive Period

In 1950, a report appeared that first indicated that the neonatal response to stress deviated markedly from that observed in the adult and thus created a field of inquiry that has persisted for over 4 decades. Using changes in adrenal ascorbic acid as the response to stress, Jailer reported that the neonate did not show any response to stress. By the early 1960s, Shapiro placed a formal label on this phenomena, designating it the stress nonresponsive period (SNRP). It is important to note that for the most part the database for this description was the inability of the rat pup to show significant elevations of corticosterone following stress. There was one study that received little attention at the time but did raise important questions concerning the validity of the notion of an SNRP. In that study, in addition to exposing the pup to stress and its demonstrating a lack of a corticosterone response, another group was injected with ACTH. These pups also failed to elicit a corticosterone response, which indicated that one of the factors contributing to the SNRP could be a decreased sensitivity of the adrenal to ACTH. Therefore, it was conceivable that other components of the HPA axis might be responsive to stress. The resolution of this question was dependent on the availability of relatively easy and inexpensive procedures for examining other components of the HPA axis. The methodological break through, which altered most of endocrinology and had a major impact on our understanding the ontogeny ofthe stress response, was the development of radioimmuneassay (RIA) procedures.

The initial impact of the RIA was to change the designation of this developmental period for the SNRP to the stress hyporesponsive period (SHRP). This change was a result of studies that showed a small but significant rise in corticosterone when measured by the RIA. Thus, although the response of the adrenal was reduced markedly during the SHRP, the adrenal was capable of releasing low levels of corticosterone.

It has been argued that the biological function of the SHRP is to protect the developing organism, specifically the brain, from an excess of circulating glucocorticoids. There is abundant evidence that high levels of glucocorticoids suppress neural growth; thus, under basal circumstances, it seems beneficial for the neonate to suppress the activity of the stress system. However, in situations that threaten health or even survival, an activation of the HPA system might be crucial. The HPA axis during development is unique in that it maintains a delicate balancing act between the regulation of stable levels of corticosterone and the ability to respond to systemic stimuli that may be life threatening. Thus, the neonate can more precisely discriminate between stressors that require higher-level cognitive processing and life-threatening events.

Indeed, when investigators began to examine other components of the neonate's HPA axis it became apparent that the SHRP was by no means absolute and that the different components of the axis responded differentially and were stimulus-specific. The question addressed in this article is whether the concept of the SHRP is still valid. In order to confront this question, we examine the development of several components of the HPA axis: the adrenal, the pituitary, and the brain.

Stress Hyporesponsive Period, the Adrenal, and Corticosterone

This article focuses on how the development of the different components of the HPA axis bears on the issue of the SHRP. It is generally agreed that in response to most stressors the neonate fails to elicit adrenocortical response or does so minimally. There are several features that characterize the function of the pup's adrenal. The first and most obvious characteristic of the adrenal is that basal levels of corticosterone during the SHRP are considerably lower than observed immediately following parturition and after the pup emerges from the period of adrenal quiescence, which is normally between postnatal day 4 and 14 in the rat (postnatal day 1 and 12 in the mouse). Further, numerous investigators have reported that the neonate can elicit a significant increase in plasma levels of corticosterone following a variety of challenges. However, invariably the

magnitude of the response is small compared to older pups that are outside the SHRP and, of course, to the adult. Thus, whereas the reported changes in corticosterone levels following stress in the adult can at times exceed 50 μg dl^{-1}, rarely does the infant during the SHRP reach levels that exceed 10 μg dl^{-1}. These levels are reached only under special circumstances. Thus, the ability of the neonatal adrenal to secrete corticosterone seems to be impaired markedly.

There is an important caveat in making the assumption that the reduced level of corticosterone following stress indicates a reduction in biological activity. Corticosterone exists in the circulation in two forms, bound and unbound. The large majority of corticosterone in the adult is bound to corticosteroid-binding protein (CBG) and other binding proteins. Only a small fraction exists in the free form, which is considered to be the biologically active form. Another aspect of the SHRP in rodents is the relative absence of CBG during the SHRP. Thus, although absolute values of corticosterone, which normally include both bound and unbound hormone, are very low in the absence of CBG, the fraction of corticosterone that is available in the free form for binding to corticosteroid receptors may actually be higher than is observed in the adult. Further, the clearance of corticosterone from the circulation is significantly slower in the pup and, as a consequence, corticosterone is available for a more prolonged period. Thus, the biologically active corticosterone does have a more prolonged time to exert its effects in the periphery and the brain.

Although there appear to be rate-limiting factors that act developmentally to limit the secretion of corticosterone in the neonate, evidence indicates that the adrenal is actively suppressed during the SHRP. It has been extensively documented that certain aspects of the rodent maternal behavior play an important role in regulating the neonate HPA axis. In particular, two specific components of the dam's caregiving activities seem to be critical; licking/stroking and feeding. Numerous studies have demonstrated that feeding is in part responsible for the suppression of the pup's capacity to both secrete and clear corticosterone from the circulation. Thus, removing the mother from the litter for 24 h (and thus fasting the pup) results in a significantly higher basal level and a further increase in the secretion of corticosterone following stress or administration of ACTH. The initial activation of the HPA axis occurs after between 4 and 8 h of maternal absence. This process is likely mediated by metabolic signals because replacement with glucose or blockade of ghrelin receptors during deprivation prevents HPA axis activation. We have postulated that one of the consequences of maternal deprivation is to increase the sensitivity of the adrenal to ACTH. This has been demonstrated in several ways. (1) Following deprivation, significantly lower doses of ACTH are required to induce the adrenal to secrete corticosterone. (2) Although the levels of ACTH are equivalent between deprived and nondeprived pups under certain experimental conditions, the levels of corticosterone are greater in deprived pups. (3) Studies indicate that following a mild stress (injection of isotonic saline) there is an increase in c-*fos* gene expression in the adrenal cortex of the deprived neonate, whereas the nondeprived pup exhibited almost no detectable levels of c-*fos* mRNA. If maternally deprived pups are provided with food during the period of maternal deprivation, both basal and stress levels of corticosterone no longer differ from mother-reared pups.

Stress Hyporesponsive Period and the Pituitary

The concept of an absolute SHRP regarding the response of the pituitary following stress in the neonate is much more problematic. Whether the pituitary can show an increase in ACTH in response to stress is dependent on numerous factors. Among these are the age of the neonate, the type of stress imposed, and, once again, maternal factors. The early findings concerning the stress response of the pituitary suggested that there was a deficiency in the neonate's capacity to synthesize ACTH. Thus, as a result, the pup should exhibit a reduction in the magnitude of the ACTH stress response. However, sufficient data indicate that the pituitary of the neonate does have the capacity to synthesize and release ACTH that closely resembles the adult response. What seems to discriminate the neonate from the adult is that for the pup the response of the pituitary is much more stimulus-dependent. Further, the ability to terminate the stress response is also not fully developed and does not mature until quite late in development. Perhaps the earliest demonstration that the neonate can indeed mount an ACTH response to at least some types of challenges was a study that challenged neonates with an injection of endotoxin throughout the period from birth to weaning. At all ages, the neonate exhibited a significant elevation of ACTH that, beginning day 5, was equivalent to the adult. Of interest is that, although there was a robust ACTH response, the corticosterone response was reduced markedly from day 5 and did not begin to approach adult value until approximately day 15. The difficulty with this study is that very large doses (90% lethal dose, LD90) were used. It has been reported more recently that

the administration of interleukin (IL)-1 elicited an ACTH response in pups as early as postnatal day 6. The peak of the response followed a similar time course to that of the adult, although the magnitude of the response was significantly lower early in development. The reduced response in day 6 neonates cannot, however, be interpreted as a reduction in the neonate's capacity to produce ACTH. Three hours following adrenalectomy (ADX), a robust increase in ACTH occurs as early as day 5, presumably due to the absence of a corticosterone negative feedback signal. This magnitude of the ACTH response is as great as that seen in older neonates at day 18, which are well out of the SHRP. Under some circumstances, the pup can show an even greater ACTH response early in development than later. In contrast, brief periods of maternal separation, exposure to novelty, injections of isotonic saline, and restraint for 30 min all failed to elicit an ACTH response in normally reared pups until they escaped from the SHRP.

Recent studies in the mouse have revealed that the low and suppressed ACTH response during the SHRP is largely mediated by an enhanced glucocorticoid receptor (GR)-mediated negative feedback. Treatment with the GR antagonist RU486 resulted in a pronounced increase of basal ACTH and corticosterone values in 9-day-old pups to levels greater than in the adult (see **Figure 1**). As a different study observed, because only a small effect of RU486 is observed when it is applied locally at the paraventricular nucleus (PVN), a direct GR-mediated suppression of ACTH production and release at the pituitary level seems likely, at least toward the end of the SHRP. These data also indicate that the hypoactivity of the neonatal stress system may be restricted to the pituitary-adrenal axis.

However, so far there is little known about why and how neonates discriminate between different classes of stimuli, whereas older pups, which have escaped from the SHRP, and adults appear to respond in a similar manner regardless of the stress-inducing stimulus. Several hypotheses could account for this phenomenon. First, it could simply be a matter of stimulus intensity. Thus, the neonate may be less responsive to stimuli of lower intensities and therefore require a more intense stressor to activate the neuroendocrine cascade that eventually leads to a release of ACTH. Second, it has been well documented that different stimuli activate distinct neural pathways that lead to the release of CRF and thus ACTH. It is conceivable that these neural pathways mature differentially or are activated during development distinctly differently than in adulthood.

A third factor appears to contribute to the reduced capacity of the mother-reared pup to respond to milder stress-inducing procedures. The role of mother's caregiving activities on the developing adrenal was discussed earlier. Evidence shows that maternal factors can also actively inhibit the release of ACTH. Maternally deprived mouse pups show a marked increase in ACTH secretion between 4 and 8 h following removal of the mother. Pups deprived of maternal care for 24 h showed an increase in ACTH following an injection of saline as early as day 6. The effects of maternal deprivation were even more apparent at days 9 and 12. Although in subsequent experiments the response at day 6 was not reliable, significant increases in ACTH were replicated in 9- and 12-day-old neonates. These increases are also observed following 30 min of restraint. In contrast, nondeprived pups failed to show an acute release of ACTH following stress. Whereas feeding was required in order to reduce the sensitivity

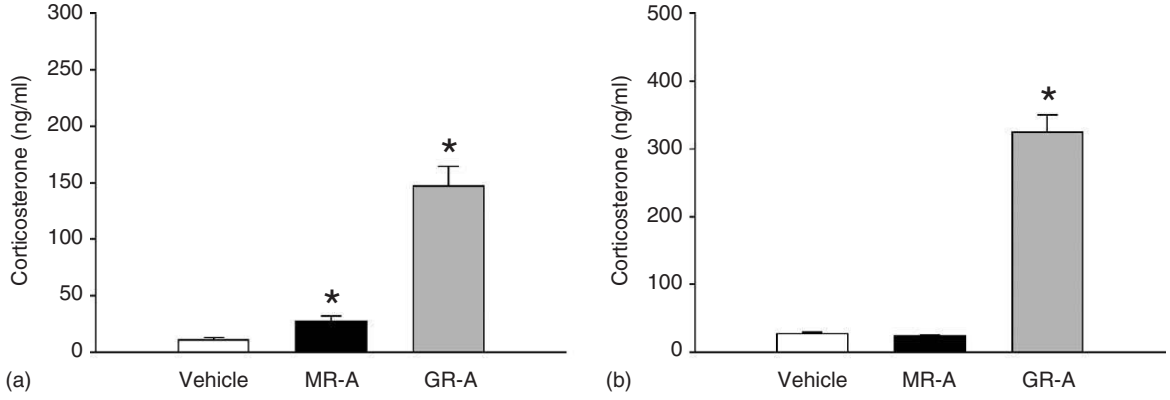


Figure 1 Nine-day-old mouse pups during the SHRP: a. plasma corticosterone; b. ACTH. Animals were injected intraperitoneally (IP) 16 and 8 h before testing with the vehicle (polyethylene glycol, $n = 10$), the MR antagonist (MR-A) RU28318 ($n = 15$), or the GR antagonist (GR-A) RU38486 ($n = 15$). Data represent mean ± SEM. * Significant from vehicle-treated animals, $p < 0.05$. Adapted from Schmidt et al., Endocrinology 146(3):1458–1464, with permission. Copyright 2005, The Endocrine Society.

of the adrenal, anogenital stroking can reverse the increased sensitivity of the HPA axis following deprivation. Thus, different components of the mother's behavior appear to be involved in regulating different components of the endocrine stress response. These data suggest that the dam's behavior was actively inhibiting part of the neuroendocrine cascade that ultimately results in the peripheral endocrine responses to stress. Thus the capacity to respond even to mild stimuli is present early in development but is observable only if the maternal inhibitory factors are not present. It is important to note, however, that although the pup can be induced to show an endocrine response to stress during the SHRP, maternal inhibition is not the only rate-limiting factor. If we examine the body of data on the ACTH responses in pups, it emerges that, even when the infant responds to mild stress during the SHRP, the magnitude of the response is always considerably lower in the SHRP than that of the older pups (day 18) and adults.

Stress Hyporesponsive Period and the Brain

The brain is clearly a stress-sensitive organ. It is not the purpose of this article to review all the changes in the brain that have been shown to occur in response to stress in the neonate. For this review, the focus is on the specific neural markers that have a direct effect on the regulation of the pituitary. Several of the early studies reported that during postnatal development there is a deficiency in the neonate's capacity to increase gene transcription and peptide expression of ACTH secretagogs, even though there is a striking increase in ACTH release. However, recent data indicate that this is not completely correct. A mild stressor in the form of a saline injection results in a fast and significant increase of CRF mRNA expression in the PVN in pups during the SHRP within 30 min rather than the 4 h used to demonstrate CRF mRNA in adult animals. There is also clear evidence that CRF is involved in the stress responses during the SHRP. The administration of a CRF antagonist blocked the ACTH response in neonates following exposure to an endotoxin. In addition, recent studies with CRF receptor 1-deficient mice demonstrated that a functioning CRF system at the pituitary level is essential for deprivation-induced changes in corticosterone secretion.

Corticosteroid Feedback

Negative feedback has two modes of operation. First, the proactive mode involves the maintenance of basal levels of HPA activity and is, in the adult, predominantly mediated by the high-affinity mineralocorticoid receptors (MRs) for corticosterone in higher brain regions. Second, the reactive mode facilitates the termination of stress-induced HPA activity, which also involves the lower-affinity GRs localized in higher brain areas as well as the PVN and the pituitary corticotrophs. Both feedback modes are operative during the SHRP. Early studies reported a pronounced increase of circulating ACTH in response to ADX in neonates, a process that was attributed to the MR. During the SHRP, the MR are highly expressed in the hippocampus, have adultlike properties and, as observed in the adult, have low amounts of corticosterone mainly occupy MRs. However, in a recent study, the treatment of neonatal mice with the MR antagonist RU28318 resulted in only a small increase in circulating corticosterone, indicating a modulatory role of the MRs. In contrast, the GR antagonist RU486 caused a profound increase of ACTH and corticosterone to adultlike levels, indicating that GRs play a predominant role in suppressing the HPA axis activity in the neonate. Because a local blockade of the GRs in the PVN of a neonatal rat resulted only in minor changes of corticosterone secretion, the most likely site of GR-mediated negative feedback in the neonate under basal conditions is in the pituitary.

The reactive feedback, probably mediated by GRs in the brain, seems to be a late-developing process because during neonatal life the stress-induced ACTH levels are not terminated as efficiently as in adulthood. Low levels of GRs are present in the brain around birth; the levels rise slowly and do not achieve adult levels before 4 weeks of age. The affinity of GRs is higher perinatally than at later ages. The microdistribution of GRs changes dramatically in some regions during the first week of life. In some regions of the hypothalamus, GR expression is very dense, but it disappears close to detection levels in later life. Other regions, such as the dentate gyrus of the hippocampus, do not express GRs until day 12.

Twenty-four hours of maternal deprivation impairs the reactive (rather than the proactive) control of HPA activity in the neonates. This was demonstrated in a study design using ADX animals and corticosterone substitution either at the time of ADX or 3 h after ADX, when the removal of the adrenals has triggered a profound ACTH response. The latter condition tests reactive feedback, which appeared during maternal deprivation to be resistant to corticosterone. Corticosterone replacement at the time of ADX, which tests the efficacy of proactive feedback, was not affected by maternal deprivation. The resistance to corticosterone could be due to two interacting

processes. The first represents the hyperdrive of signals from the stress circuitry innervating the PVN neurons triggered by maternal deprivation. The second process may be related to an effect of maternal deprivation on the development of corticosteroid receptors. Indeed, maternal deprivation downregulates hippocampal MR expression as well as GR expression in discrete regions of the hippocampus and the PVN. These changes in GR could be restored if the deprived neonate received aspects of maternal care, such as stroking and feeding. Thus, the altered characteristics clearly may be involved in the impaired feedback following maternal deprivation.

Conclusion

Is there an SHRP? The concept of the SHRP in its most traditional form describes a period in development when most aspects of the HPA axis appear to be downregulated, resulting in an organism that either fails to respond or at best shows a response that is reduced markedly compared to older pups and adults. This may still be applicable to the response of the adrenal if we take total corticosterone as the index of the stress response. Therefore, the SHRP in fact represents an adrenal hyporesponse period, primarily caused by a potent corticosterone feedback signal that ensures quiescence in the face of common daily stressors. The existing information on the ontogeny of the HPA axis does not support, therefore, the concept of an SHRP as it was formulated in the early 1970s.

Acknowledgments

This work was supported by a grant from the National Institutes of Mental Health, MH-45006 (to S.L.), a NATO collaborative research grant (CRG 97.0477), the Netherlands Organization for Scientific Research (NWO), 554–545 and the Royal Netherlands Academy of Arts and Sciences.

See Also the Following Articles

Corticosteroid-Binding Globulin (Transcortin); Corticotropin Releasing Factor (CRF); Glucocorticoid Negative Feedback.

Further Reading

Bohn, M. C. (1984). Glucocorticoid-induced teratologies of the nervous system. In: Yanai, J. (ed.) *Neurobehavioral teratology*, pp. 365–387. New York: Elsevier.

de Kloet, E. R., Sibug, R. M., Helmerhorst, F. M., et al. (2005). Stress, genes and the mechanism of programming the brain for later life. *Neuroscience and Biobehavioral Reviews* **29**(2), 271–281.

de Kloet, E. R., Rosenfeld, P., Van Eekelen, J. A. M., Sutanto, W. and Levine, S. (1988). Stress, glucocorticoids and development. *Progress in Brain Research* **73**, 101–120.

Levine, S. (1994). The ontogeny of the hypothalamic-pituitary-adrenal axis: the influence of maternal factors. *New York Academy of Sciences* **740**, 275–288.

Levine, S., Glick, D. and Nakane, P. K. (1976). Adrenal and plasma corticosterone and vitamin A in rat adrenal glands during postnatal development. *Endocrinology* **118**, 1177–1179.

Meaney, M. J., Dioro, J., Francis, D., et al. (1994). Environmental regulation of the development of glucocorticoid receptor systems in the rat forebrain: the role of serotonin. *New York Academy of Sciences* **740**, 260–274.

Rosenfeld, P., Suchecki, D. and Levine, S. (1992). Multifactorial regulation of the hypothalamic-pituitary-adrenal axis during development. *Neuroscience and Biobehavioral Reviews* **16**, 553–568.

Rosenfeld, P., Van Eekelen, J. A. M., Levine, S., et al. (1993). Ontogeny of corticosteroid receptors in the brain. *Cellular and Molecular Neurobiology* **13**, 295–321.

Sapolsky, R. M. and Meaney, M. (1986). Maturation of the adrenocortical stress response: neuroendocrine control mechanism and the stress hyporesponsive period. *Brain Research Reviews* **11**, 65–76.

Schmidt, M., Levine, S., Oitzl, M. S., et al. (2005). Glucocorticoid receptor blockade disinhibits pituitary-adrenal activity during the stress hyporesponsive period of the mouse. *Endocrinology* **146**(3), 1458–1464.

Stress in University Students

B Andrews and J Hejdenberg
Royal Holloway University of London, Egham, UK

© 2007 Elsevier Inc. All rights reserved.

Conceptual Issues in Student Stress Research
The Impact of Commonly Identified Student Stressors
Implications for Intervention

Glossary

Cross-sectional research	Investigations in which all factors are assessed at one time.
Homesickness	Distress from longing for home during one's absence from it.
Life-stage transition	Moving from one life stage to another (e.g., from childhood to young adulthood) with attendant changes (e.g., from being materially dependent on others to becoming more independent).
Prospective/ longitudinal research	Investigations in which factors are assessed at more than one time; usually carried out to identify factors that might predict future outcomes.

Pressures facing university students have been a cause of concern because of their impact both on mental health and academic performance. In recent years, student mental health services in the United Kingdom and United States have reported increases in the severity of the presenting problems reported by students. There is also evidence from studies in the United Kingdom that university students report more mental health symptoms than age-matched controls, although there is a lack of information on diagnosable mental disorders in students internationally. In some countries, apparent increases in student distress have been linked to decreases in state funding for students and widening student participation; widening access to higher education has resulted in more diverse student populations, thereby increasing the probability of students in the system with elevated levels of life problems. Given the paucity of information on specific sources of stress and student physical health, the focus in this article is on the stressors associated with the university experience that relate to poor mental health and academic performance.

Conceptual Issues in Student Stress Research

In research on university students, the term stress has commonly been used to describe subjective responses to an experience or situation. Homesickness and worries associated with meeting course requirements are good examples, although non academic worries are often included. The term has also been used to refer to life events and difficulties such as bereavement or financial hardship that may be considered universally threatening without the proviso in assessment that the event caused stress, upset, or worry. In the main, the literature has emphasized students' subjective worries rather than reports of actual life difficulties. Some authors have argued for the advantages of assessing stress as the amount of subjective worry associated with individual stressors. The individuality in the experience of stress, such as the availability of coping skills and/or personal vulnerability, is recognized by using such a method. The approach is useful for descriptive purposes, but research into the effects of stress on student mental health and achievement demands a more objective assessment of life problems to avoid confounding cause and effect.

The Impact of Commonly Identified Student Stressors

Before reviewing the literature, it should be noted that much of the relevant research has been carried out with narrow samples of students in vocational studies or other specialized subjects. There is therefore the problem of the generalizability of vocational samples, such as medical and nursing students, to the whole student population. The general use of cross-sectional designs, with only one data collection point, is also problematic because causal direction cannot be inferred. The relevant studies described here have therefore been selected because they represent broader populations of university students, and, whenever possible, prospective studies are described.

The sources of stress most commonly identified by students in surveys worldwide include the academic, financial, and relationship domains, although it should be noted that student reports are constrained by the areas initially defined by the investigators. Researchers have also identified the transition to university and paid term-time employment as additional sources of stress. To evaluate the existing evidence of the effects of stress on student mental health and achievement, four of these stressors (transitions, academic and financial pressures, and term-time employment) are discussed here. These have been selected because they are specifically related to the student experience and have been the main focus of investigations into the impact of stress on students.

Transition to University

Becoming a student involves major life changes for young people. Faced with new responsibilities and social roles inherent to the university experience, this transition has the potential to cause distress. The literature has focused on the negative aspects of this transition, but it can also provide the opportunity to learn new competencies and to develop new and rewarding social relationships. For example, longitudinal studies of life-stage transitions have demonstrated their positive effects on self-esteem in young adulthood. However, few longitudinal student studies have been undertaken, and studies that have followed students through the transition to university with a survey before entry are uncommon. Studies that have surveyed students before university entry have shown (usually) short-term increases in mental health symptoms after university entry, which on the face of it appears to be indicative of the stress associated with this experience. In one important study of homesickness, Fisher and Hood measured psychological disturbance in students 3 months before university entry and again 6 weeks into their first term. The levels of psychological disturbance, which included depressive and obsessional symptoms and absent-mindedness, increased in the period after the transition to university. The increases were most apparent in students reporting homesickness, although these students had higher symptom levels to start with. Using an objective measure of residence change, there was no difference in psychological disturbance between students still living with parents and those living on campus. This research demonstrated that, at least in the short term, entering university has consequences in terms of emotional disturbance and cognitive impairment regardless of residence status. This suggests that factors other than the transition of moving from home contribute to student distress.

A subsequent longitudinal study by Andrews and Wilding assessed anxiety and depressive symptoms in students and showed a significant increase in symptoms from 1 month before college entry to midcourse. However, when they looked categorically at the rates of high depression and anxiety in the students, a rather more complex picture emerged. Of the students who were relatively symptom-free before entry, 29% had developed high levels of depression or anxiety symptoms (subsequently explained by factors such as financial hardship and relationship problems). In contrast, approximately one-third of those with high levels of depressive or anxious symptoms before entry were relatively symptom-free by midcourse. This indicates that transitions at this life stage should not be viewed as exclusively negative. There is also a need to consider the more positive aspects of the university experience and the ways in which they may act to increase student well-being.

Academic Pressures

The university years are a time of continuous assessments for students in most countries, and as such they can put a considerable amount of pressure to achieve on even the hardiest of students. As already discussed, one of the problems in assessing the impact of the stress associated with academic work is disentangling cause from effect. Although academic pressures may have a negative effect on emotional well-being, anxiety and depression can affect academic work through symptoms such as high arousal and lack of motivation and concentration. In the absence of in-depth enquiries in this area, it is usually not possible to ascertain whether academic stress is an antecedent or consequence of mental health problems. A more useful line of enquiry into the impact of academic study *per se* is to consider specific academic events such as taking exams. This experience is recognized by leading life events researchers as a universally threatening experience that can increase the likelihood of physical and mental disorders. The effect of studying for and taking examinations has been demonstrated in a recent longitudinal study of undergraduates carried out by Surtees and colleagues at the University of Cambridge in England. Levels of anxiety and depressive symptoms peaked during examination periods, although this suggests that the negative impact is likely to be short-lived for the majority who pass successfully.

Financial Problems

Most young people have not managed their own finances before starting university, and this can often come as a rude shock. Financial hardship is a common problem facing students, and one that is expected to worsen in countries where state funding is being reduced in favor of students paying their own fees and living costs. It is therefore of particular concern that relevant studies have consistently indicated the adverse effects of financial difficulties on student mental health. There is also emerging evidence of the negative impact financial problems and their attendant consequences can have on academic achievement. Much of this research has been carried out in the United Kingdom, where there have been rapid changes to student funding in the past few years.

To address the problems of causality evident in cross-sectional research, Andrews and Wilding recently carried out a longitudinal study of over 350 undergraduate students in different university faculties to consider the impact of different types of stressful life experiences, including financial difficulties. The main

experiences assessed included major financial difficulties, serious personal illness or injury, serious illness and death of close others and serious relationship problems and separations. Of the various adverse experiences, only financial difficulties (experienced by 21%) and personal illness or injury (experienced by 6%) contributed significantly to increased levels of depression midcourse, after adjusting for depressive and anxiety symptoms present before university entry. This initial evidence suggests that the direction of causality is from financial problems to depression. The odds of becoming depressed were three times higher for students with financial problems. This study also considered the impact of stressful life experiences on academic attainment and found that the only adverse experience that predicted poor exam performance was financial difficulties. Depression appeared to mediate this relationship when prior academic attainment was controlled.

Paid Term-Time Employment

An additional stressor related to financial hardship is that of paid work. Large-scale studies have shown that students reporting difficulties in keeping up with bills or having financial problems are significantly more likely to undertake term-time employment. Being obliged to work in addition to full-time study is a further pressure on students. So far, research in this area has focused on the effects of term-time work on academic performance rather than on mental health. Large-scale surveys of university students on both sides of the Atlantic have found that academic performance decreases as the hours of paid work increase. There is evidence from student research in the United States by Pascarella and colleagues that paid term-time work can impede the basic cognitive processes necessary for academic achievement such as reading comprehension and critical thinking. Subsequent studies have found that students who work longer hours during term-time achieve lower end-of-year exam averages than other students. In a recent study, this effect still held after relevant demographic factors and prior academic performance were taken into account.

Implications for Intervention

Student counseling services are a common feature of university life, along with other educational support services and pastoral care from personal tutors and advisors. However, to our knowledge there is little available information on the systematic evaluation of the effectiveness of such services. Some real-life difficulties that cause students distress may not be amenable to the kinds of services that counselors provide, particularly when they involve financial crises and the pressures of juggling the need to earn money with academic demands. Money management courses would appear to be relevant, but more needs to be known about their availability and effectiveness. Counseling services may be more relevant in dealing with stressors of a more intrapsychic nature. Most services deal with issues of attachment and homesickness, and many provide support in managing the stress experienced by students around exam times. Given the available evidence of the adverse effects of these common problems on student mental health and academic attainment, there is an obvious need for well-controlled trials to evaluate how well university support services deal with different types of student stressors.

See Also the Following Articles

Economic Factors and Stress; Psychosocial Factors and Stress.

Further Reading

Andrews, B. and Wilding, J. M. (2004). The relation of depression and anxiety to life-stress and achievement in students. *British Journal of Psychology* **95**, 509–521.

Fisher, S. (1994). *Stress in academic life: the mental assembly line*. Buckingham, UK: Open University Press.

Grant, A. (2002). Identifying students' concerns taking a whole institutional approach. In: Stanley, N. & Manthorpe, J. (eds.) *Students' mental health needs: problems and responses*. London: Jessica Kingsley.

Pascarella, E. T., Edison, M. I., Amaury, N., et al. (1998). Does work inhibit cognitive development during college? *Educational Evaluation and Policy Analysis* **20**, 75–93.

Royal College of Psychiatrists (2003). *The mental health of students in higher education*. Royal College of Psychiatrists Council report CR112, London: RCP.

Stewart-Brown, S., Evans, J., Patterson, J., et al. (2000). The health of students in institutes of higher education: an important and neglected public health problem? *Journal of Public Health Medicine* **22**, 492–499.

Surtees, P. G., Wainwright, N. W. J. and Pharoah, P. D. P. (2002). Psychosocial factors and sex differences in high academic attainment at Cambridge University. *Oxford Review of Education* **28**, 21–38.

Stress Induced Anovulation

S L Berga and T L Loucks
Emory University School of Medicine, Atlanta, GA, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction
Neuroendocrine Mechanisms Linking Cognition, Mood, Behavior, and Gonadotropin-Releasing Hormone Drive
Pathogenesis of Stress-Induced Anovulation
Behavioral, Nutritional, and Metabolic Influences on the Reproductive Axis
Synergism among Stressors
Treatment Considerations
Summary

Glossary

Allostasis	A sustained adjustment that promotes survival in response to chronic challenge; may increase long-term health burden.
Anovulation	The cessation of ovulation.
Cognitive-behavior therapy	A form of psychoeducation or talk therapy that addresses nonadaptive thoughts, cognitions, attitudes, and conceptualizations.
Eumenorrhea	A normal pattern of menstrual bleeding; can occur in the absence of ovulation.
Gonadotropin-releasing hormone (GnRH)	Decapeptide produced and released in a pulsatile mannner from hypothalamic neurons; stimulates the release of pituitary lutenizing hormone (LH) and follicle stimulating hormone (FSH).
Hypothalamic hypogonadism	A condition characterized by suppression of GnRH drive and, in turn, pituitary LH and FSH, resulting in reduced or absent gonadal steroidogenesis and gametogenesis.
Hypothalamic hypothyroidism	An allostatic adjustment in thyroidal function that conserves energy expenditure and is characterized by relatively normal levels of pituitary thyroid stimulating hormone (TSH) in the presence of suppressed levels of thyroid hormones triiodothyronine (T3) and thyroxine (T4).
Limbic-hypothalamic-pituitary-adrenal axis	Neuroendocrine circuit that mediates stress perception and stress responses.
Metabolism	Biochemical processes that regulate energy expenditure and storage.
Secondary amenorrhea	Cessation of menses in a woman of reproductive age who was previously eumenorrheic.
Stress	Physical and psychological stimuli that challenge the status quo and elicit homeostatic or allostatic behavioral responses.

Introduction

A wealth of scientific and clinical evidence supports the notion that stress causes reproductive compromise and increases health burden. Neuroendocrine, metabolic, and behavioral responses to acute stress represent transient homeostatic adaptations that promote survival in the face of perceived challenge; chronic stress elicits allostatic (sustained) adjustments that also promote survival but at greater health cost (Figure 1). Reproductive compromise impedes population replenishment and increases acute and chronic health burden in women, men, and children due to metabolic dysfunction, obesity, diseases of aging, preterm delivery, birth defects, costly and risky infertility therapies, and, through epigenetic mechanisms, it imprints the next generation. Stress is the most common and most commonly underappreciated cause of reproductive dysfunction (Table 1). Stress-induced anovulation (SIA), often termed functional hypothalamic amenorrhea (FHA) or functional hypothalamic chronic anovulation, causes infertility and increases acute and chronic health burden.

Behaviors that chronically activate the limbic-hypothalamic-pituitary-adrenal (LHPA) axis and/or chronically suppress the hypothalamic-pituitary-thyroidal (HPT) axis compromise the hypothalamic-pituitary-gonadal (HPG) axis in both women and men by reducing hypothalamic gonadotropic-releasing hormone (GnRH) drive. Individuals with functional (not due to defined organic conditions) forms of hypothalamic hypogonadism typically engage in a combination of behaviors in response to psychogenic stress that concomitantly induce intermittent or chronic energy imbalance. Further, chronic adrenal activation may increase the energetic cost of common activities such as running. In women, functional hypothalamic hypogonadism exists on a spectrum that includes polymenorrhea (menstrual interval < 24 days), eumenorrhea with reduced luteal progesterone secretion, and anovulatory eumenorrhea, oligomenorrhea (menstrual interval 735 days), or amenorrhea (absence of menstruation for >3–6 months). In men, oligoasthenospermia (very low sperm count with low motility and abnormal morphology) may result, but this is typically not clinically evident unless fertility is being sought or severe testosterone deficiency results in muscle wasting or other phenotypic alterations.

This article focuses primarily on SIA/FHA in women. SIA/FHA typically results from psychogenic stress coupled with a mild energy imbalance and represents an allostatic adaptation, that is, a stable

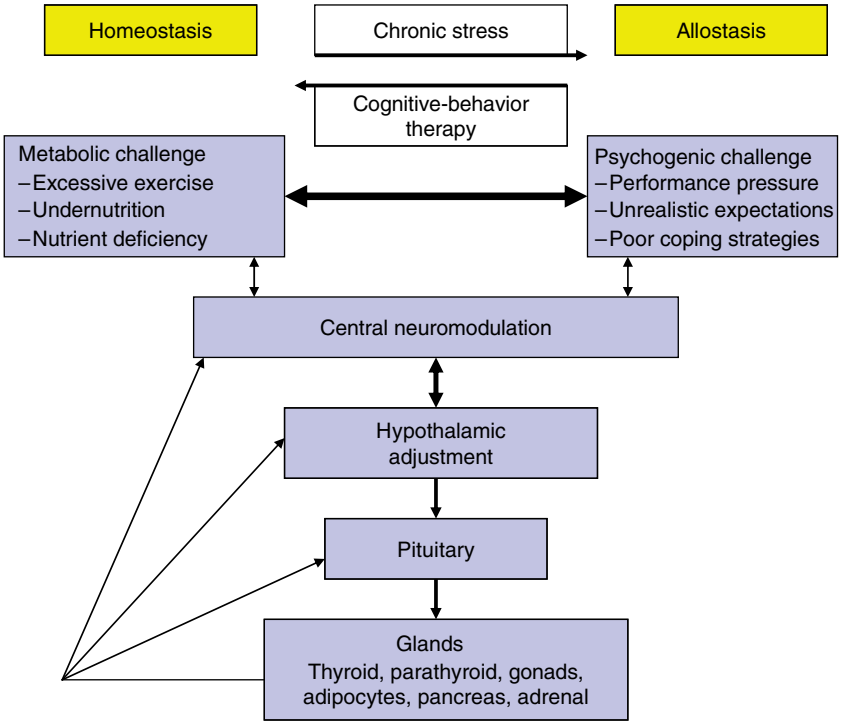


Figure 1 Conceptualization of synergism between metabolic and psychogenic stress in the pathogenesis of stress-induced anovulation.

Table 1 Common causes of anovulation/amenorrhea[a]

	Percentage
Stress-induced anovulation/ functional hypothalamic amenorrhea	34
Hyperandrogenism/PCOS	29
Hyperprolactinemia	13
Premature menopause	12
Asherman's syndrome	5
Other	7

[a]Based on data presented in Reindollar, R. H., Novak, M., Tho, S. P. T., et al. (1986), Adult-onset amenorrhea: a study of 262 patients, *American Journal of Obstetrics and Gynecology* **155**, 531–543. PCOS, polycystic ovarian syndrome.

change in behaviors and secretory patterns that promotes acute survival but at some health cost. SIA/FHA affects roughly 5% of women of reproductive age; less severe forms of hypothalamic hypogonadism are more common and are less clinically evident (Figure 2).

Stress-induced anovulation is theoretically reversible, but reproductive recovery appears to depend on the restoration of eucortisolemia and at least partial recovery from functional hypothalamic hypothyroidism. Hormone replacement strategies have limited benefit for women with SIA/FHA because they do not promote recovery from allostatic endocrine adjustments in the adrenal and thyroidal axes. Indeed, the rationale for the use of sex steroid replacement is based on the erroneous assumption that functional forms of hypothalamic hypogonadism represent only or primarily a loss of sex steroid exposure due to reduced GnRH drive. Further, the replacement of sex hormones masks deficits that accrue from chronically altered adrenal and thyroidal functions. Long-term deleterious consequences of SIA/FHA probably include an increased risk of cardiovascular disease, osteoporosis, depression, other psychiatric conditions, dementia, and neurodevelopmental compromise in offspring. Although fertility can be restored with exogenous administration of gonadotropins or pulsatile GnRH, fertility management alone does not engender adrenal and thyroidal recovery. Pregnancy in the face of ongoing psychogenic stress and metabolic imbalance may increase the likelihood of poor obstetrical, fetal, or neonatal outcomes. In contrast, behavioral and psychological interventions that address problematic behaviors and attitudes have the potential to facilitate reproductive recovery along with adrenal and thyroidal recovery. In short, full endocrine recovery offers better individual, maternal, and child health.

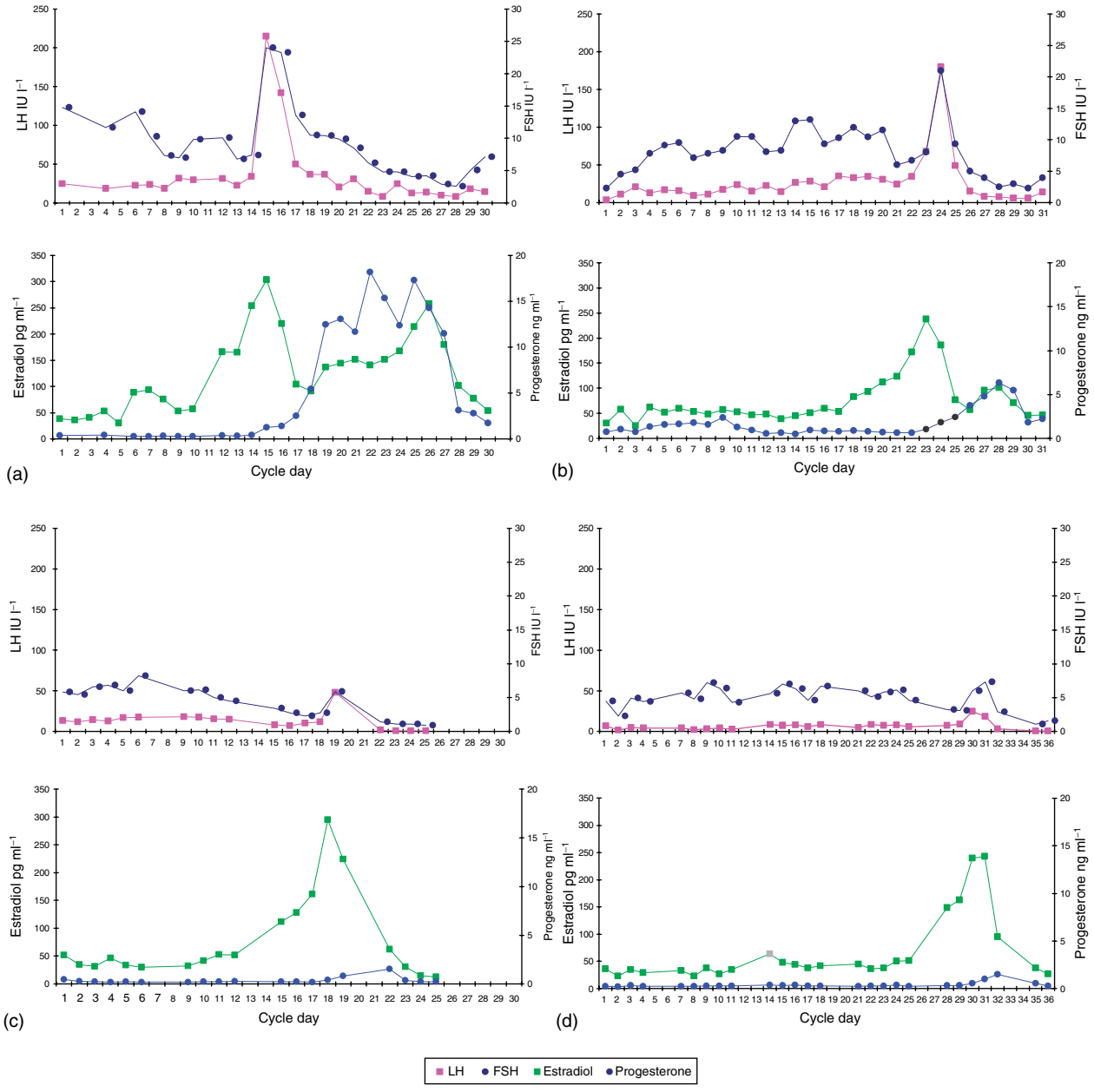


Figure 2 Continuum of ovarian function in women with preserved menstrual cyclicity. a, Ovulatory eumenorrhea; b, luteal insufficiency; c–d, eumenorrheic anovulation. Pituitary gonadotropin (LH and FSH) and ovarian sex steroid (estradiol and progesterone) levels were obtained daily for one cycle. FSH, follicle-stimulating hormone; LH, luteinizing hormone. Note that FSH > LH, but FSH < 20 IU/L, thus the cause of hypogonadism is hypothalamic rather than reduced ovarian reserve.

Neuroendocrine Mechanisms Linking Cognition, Mood, Behavior, and Gonadotropin-Releasing Hormone Drive

Our conceptualization of the bidirectional interaction between behavior and gonadal function has been refined by scientific insights into the mechanisms mediating neuroendocrine responses to thoughts, feelings, and behaviors.

The proximate cause of hypothalamic forms of hypogonadism, including SIA/FHA, is reduced hypothalamic GnRH input. Gonadal function depends directly on secretion from the hypothalamus of GnRH as pulses. Declines in pulsatile GnRH secretion reduce pituitary secretion of luteinizing hormone (LH) and follicle-stimulating hormone (FSH) and wholly or partially compromise folliculogenesis.

Decreased GnRH drive is a common cause of anovulation and amenorrhea. Decrements in central GnRH–LH/FSH drive exist on a continuum, however, and may vary from day to day. Because of this potential variability in GnRH drive, ovarian compromise exists on a spectrum and may present as polymenorrhea, eumenorrhea, oligomenorrhea, or amenorrhea. Clinically, the decreased ovarian function can be occult or obvious. In men, decreased central GnRH drive may cause oligoasthenozoospermia. Typically, gonadal compromise in men is clinically occult unless fertility is sought and the compromise is sufficiently significant to cause infertility. However, severe hypothalamic hypogonadism in men may present as decreased libido, diminished muscle mass, or altered hair (beard) growth.

The most common cause of reduced GnRH drive is functional; that it, it is not due to identifiable organic causes such as hypothalamic tumors or pituitary adenomas. Functional hypothalamic hypogonadism can be defined as a common and theoretically reversible form of gonadal compromise in which psychophysiological responses to life events activate central neural networks and thereby alter glandular secretion and metabolism.

The mechanisms mediating the disruption of central GnRH drive are poorly understood and complex. GnRH neurons are diffusely distributed in the medial basal hypothalamus. Most, but not all, GnRH axons project to the median eminence, allowing pulses of GnRH to be released into the portal vasculature. GnRH neurons communicate with one another via synapses. Although GnRH neurons are endogenously pulsatile, their activity must be synchronized by GnRH-to-GnRH synapses for the GnRH bolus released into the portal vasculature to be of sufficient magnitude to trigger pituitary release of LH and FSH. GnRH neurons receive synapses from neurons that contain GnRH, the endogenous opioid peptide β-endorphin, neuropeptide Y (NPY), and catecholamines. Other factors that modulate the frequency or activity of the GnRH pulse generator are thought to exert their effects indirectly by acting through the neuronal systems that have direct synaptic connections or by interacting with glial cells that interpose between synapses. For instance, progesterone appears to slow the frequency of pulsatile GnRH release by increasing hypothalamic opioidergic tone, but it also may cause glial cells to envelope the GnRH neurons and to reduce thereby the GnRH-to-GnRH connections. Peripheral substances may gain access to the GnRH neuronal network via specialized neurovascular cells that line the fenestrated blood–brain barrier at the level of the hypothalamus and median eminence, so the modulation of GnRH secretion is not confined to central factors. The brain–gut axis appears to communicate the metabolic status of an individual to the hypothalamic GnRH neurons through multiple mechanisms. A partial list of putative neuromodulators of GnRH drive is shown in Table 2.

Table 2 Putative modulators of GnRH drive[a]

CRH	Metabolic signals
Opioids	Sex steroids
Adrenergic	Gonadal peptides
GABA	Growth factors
Dopamine	Glial cells
Serotonin	GnRH
Immune	Other

[a]CRH, corticotropin releasing hormone; GABA, γ-aminobutyric acid; GnRH, gonadotropin-releasing hormone.

Progress has been made in quantifying endocrine patterns for research purposes. However, in humans, GnRH pulsatile secretion can be inferred only from the pattern of LH secretion in the circulation. Blood samples must be obtained via an indwelling intravenous catheter at intervals of 10–15 min for durations of 12–24 h. Even so, inherent limitations exist in estimating GnRH secretion from peripheral LH patterns. The tedious nature of quantifying central GnRH drive explains why this is not done routinely during clinical evaluation. Technical limitations also plague the recognition and quantification of stress and metabolic status. The accuracy of psychometric inventories for assessing and quantifying stress, mood, and cognitive patterns is inherently constrained by reporting biases, whereas endocrine or biophysical indices of stress and metabolic status are technically cumbersome and expensive to collect. Neuroimaging techniques have afforded a new window into the neurochemistry and neuroanatomy of behaviors and thoughts, but these techniques have not yet been used to understand the pathogenesis of stress-induced reproductive compromise.

Although innumerable animal studies have demonstrated that activation of the hypothalamic-pituitary-adrenal (HPA) axis by a variety of stressful paradigms induces reproductive compromise, only a few studies have elucidated the mechanisms mediating the disruption of GnRH drive. Direct evidence exists for corticotropin releasing hormone (CRH), β-endorphin, dopamine, and arginine vasopressin. Recently, NPY has been implicated as serving a neuromodulatory link between metabolic deficits induced by diet and exercise and reduced GnRH drive. A key inhibitory neurotransmitter system in the brain uses γ-aminobutyric acid (GABA). GABA opens the same potassium channels in neurons of the mediobasal

hypothalamus as μ-opioid receptor agonists such as β-endorphin. GABA inhibited GnRH gene expression in rats and suppressed pubertal GnRH increase in juvenile female rhesus monkeys, but the role of GABA in human hypothalamic hypogonadism remains unclear. Serotonin neurons may also play a role in modulating GnRH drive.

Pathogenesis of Stress-Induced Anovulation

The best biochemical evidence supporting the concept that stress impairs ovarian function in women is the consistent demonstration that women with hypogonadotropic hypogonadism not due to defined organic conditions have higher cortisol levels than eumenorrheic, ovulatory women. There is direct evidence that this relationship holds in women with athletic amenorrhea as well. In a study by Loucks et al., an inverse relationship was observed between the degree of ovarian compromise and cortisol levels. When compared with eumenorrheic but sedentary women, eumenorrheic athletes had less luteal progesterone secretion, as evidenced by lower urinary levels of pregnanediol-glucuronide, fewer LH pulses in a day, and higher cortisol levels. Furthermore, amenorrheic athletes that were anovulatory had the fewest LH pulses in a day and the highest cortisol levels despite comparable levels of exertion and fitness. An inverse relationship holds between adrenal activation (independent of the life events or behaviors that initiate or sustain this activation) and suppression of the hypothalamic GnRH drive to the ovary, as evidenced by marked reduction in LH pulse frequency.

Other hypothalamic outputs also are altered in women with SIA/FHA. Given the neuroanatomical integration of the hypothalamus, this is predictable. The purpose of the hypothalamus is to generate an endocrine action plan to preserve the organism in the face of challenge. Part of the action plan involves metabolic mobilization. However, metabolic mobilization involves more than an increase in cortisol secretion. In SIA/FHA, the thyroid axis differs from that of eumenorrheic women in that thyrotropin (TSH) is not increased in response to decrements in thyronine and thyroxine, which indicates an altered hypothalamic set point akin to what is seen in hospitalized patients who develop what is referred to as sick euthyroid syndrome. In athletic women, a similar alteration in the thyroid axis was seen only in those who had compromised ovarian function. The secretory patterns of growth hormone, prolactin, and melatonin also differed from those of eumenorrheic women. The constellation of neuroendocrine aberrations that accompany SIA/FHA strongly suggests that central neurotransmission has been chronically altered to forge an allostatic state. The aim of allostatic adjustments is to allow the individual to cope with chronic as opposed to acute challenge. In this context, then, the hypothalamus links the external environment, the internal milieu, and gonadal function.

The process of recovering from the allostatic state of SIA/FHA is also interesting but less well documented. We showed that women in the process of recovering from functional hypothalamic hypogonadism displayed cortisol levels identical to those of eumenorrheic women before there was complete recovery of GnRH drive. Further, a marked increase in thyrotropin-stimulating hormone (TSH) drive occurred as a prelude to increases in thyroxine and thyronine. We recently reported that women with SIA/FHA treated with cognitive-behavior therapy (CBT) displayed a return of ovarian function and ovulation and reduced cortisol levels but had only partial recovery from hypothalamic hypothyroidism and no weight gain. These data, coupled with our psychometric data showing that women with SIA/FHA differ from eumenorrheic women (EW) in their attitudes toward eating and desire for thinness, raise the possibility that intermittent or mild energy imbalance may persist after CBT-induced reproductive recovery. In short, hypothalamic recovery involves a set of interlinked readjustments, including LHPA restoration, return of GnRH pulsatility, and at least partial remission of hypothalamic hypothyroidism, but the duration and sequence of recovery is just now beginning to be described and the role of metabolic factors as mediators of recovery remains unclear.

The characteristic hypothalamic alterations associated with SIA/FHA only become problematic when ongoing challenges elicit a chronic rather than acute response. The long-term consequences of persistent HPA activation have been studied in animal models and hippocampal neuron loss has been documented. Stress appears to increase the risk of dementia and other neurodegenerative disorders. Other long-term sequelae of persistent metabolic mobilization are largely unknown, but there is no reason to assume that such a process is benign. For instance, recent data obtained in women with weight-restored anorexia nervosa who remained amenorrheic indicated that exogenous sex steroid replacement was unable to stimulate appropriate bone accretion. The investigators speculated that the ineffectiveness of hormone exposure was due to ongoing metabolic derangements such as increased cortisol exposure, altered growth hormone action, or hypothalamic hypothyroidism. Because of the concomitant endocrine and metabolic disturbances, hypothalamic hypogonadism must be regarded as a condition deserving

clinical attention even when fertility is not an immediate goal.

The central neuromodulators responsible for the initiation and maintenance of the disruption of GnRH are difficult to identify in humans. First, the factors that initiate allostasis may differ from those that maintain allostasis. To study initiating factors, we would need to intensely monitor populations at risk for the development of hypothalamic hypogonadism or try to induce hypothalamic hypogonadism in a nonhuman primate model. Once a chronic hypogonadal state had been reached, it theoretically would be possible to identify the agents that maintain this disruption by administering antagonists that cross the brain–blood barrier, by performing lumbar punctures and obtaining cerebrospinal fluid (CSF), or by performing neuroimaging studies with an appropriate ligand. To date, efforts to identify these neuromodulators in humans have yielded inconsistent results. Thus, naloxone, an opioidergic blocker, increased LH pulse frequency or levels in some, but not all, women with SIA/FHA. Also, the infusion of metoclopramide, a dopamine receptor blocker, to women with FHA accelerated LH pulse frequency, whereas that of eumenorrheic women remained constant. These data suggest that there may be dopaminergic as well as opioidergic inhibition of GnRH drive. To explore the hypothesis that the reduction in GnRH drive was maintained by CRH, vasopressin, β-endorphin, or a combination of these factors, we performed lumbar punctures to obtain rostral CSF in women with SIA/FHA and those with eumenorrhea. This approach revealed increased CRH in the CSF of women with depression, but we found that CRH levels were identical in women with SIA/FHA and eumenorrhea. Vasopressin levels were similar, but, surprisingly, β-endorphin levels were lower in women with SIA/FHA. Further, we evaluated free cortisol in the CSF of these same women and found that these levels were approximately 30% higher in those women with stress-induced anovulation. In aggregate, these findings (Figure 3) support the notion that the preservation of CSF CRH levels in the presence of elevated CSF cortisol reflects chronic stress-induced (allostatic) resistance to negative feedback suppression by cortisol.

Other putative central neurotransmitters that may contribute to the initiation and maintenance of SIA/FHA are the serotonergic and GABAergic systems. Several lines of evidence suggest that serotonergic function gates stress reactivity and metabolism. First, we showed in our monkey model of SIA/FHA that stress-sensitive monkeys had larger cortisol elevations and reduced prolactin responses to fenfluramine, a serotonergic agonist. These data were interpreted as indicating that central serotonergic

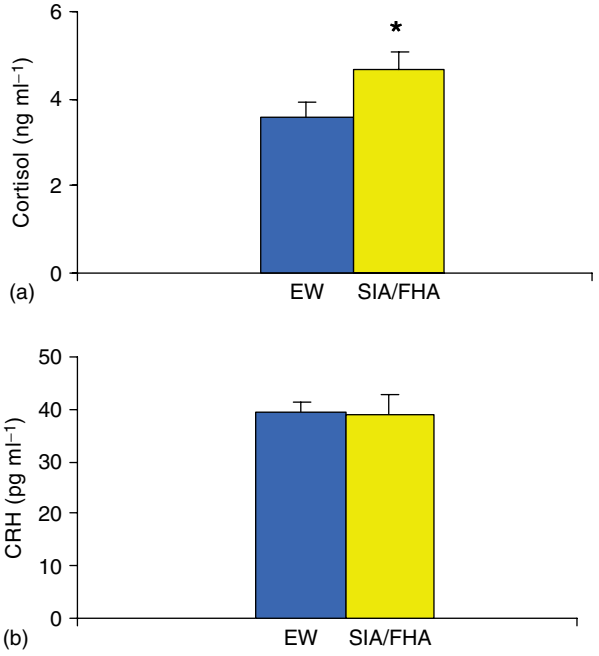


Figure 3 Mean ± SE cerebrospinal fluid hormone levels in eumenorrheic women (EW) and women with SIA/FHA, demonstrating allostatic resistance to cortisol feedback suppression of CRH. a, Cortisol levels; b, CRH levels. CRH, corticotropin releasing hormone; SIA/FHA, stress-induced anovulation/functional hypothalamic amenorrhea. From Brundu, B., Loucks, T. L., Adler, L. J., et al. (2006), Increased cortisol in the cerebrospinal fluid of women with functional hypothalamic amenorrhea, *Journal of Clinical Endocrinology and Metabolism* **91**, 1561–1565.

tone was reduced in stress-sensitive as contrasted with stress-resilient monkeys. Second, serotonergic neurons, which express leptin receptors, terminate on GnRH neurons. Serotonin has satiety effects similar to those of leptin and may modulate the response to orexigenic and anorectic signals. Moenter's lab exposed fasting female mice to saline, leptin, or fluoxetine, a selective serotonin reuptake inhibitor (SSRI) antidepressant, and followed estrus cycle length. Leptin and fluoxetine prevented fasting-induced cycle lengthening, whereas the serotonin receptor antagonist metergoline blocked cycle rescue by fluoxetine. Treating leptin-deficient *ob/ob* and leptin receptor-deficient *db/db* mice with fluoxetine did not normalize body weight or rescue estrus cycles. Further, there is abundant evidence of a role for the serotonergic system in the regulation of stress reactivity and feedback inhibition of the HPA axis by cortisol. In addition, variation in transporter-facilitated uptake of serotonin has been implicated in anxiety in humans and in animal models of anxiety and stressfulness. Human serotonin reuptake transporter (SERT) gene transcription is modulated by a common

polymorphism in its upstream regulatory region. The short variant reduces the transcriptional efficiency of the SERT gene promoter, resulting in decreased SERT expression and serotonin reuptake. The short (s) variant polymorphism is associated with an increased incidence of affective disorders in humans and maladaptive behavior in monkeys coincident with increased HPA responsivity to stress. Further, the long/short genotype produces a phenotype similar to that of the short/short genotype. In humans, the SERT polymorphisms moderated the influence of stressful life events on depression in a representative birth cohort. Stress-sensitive monkeys had a lower expression of SERT mRNA and monoamine oxidase A (MAO-A) in the caudal region of the dorsal raphe nucleus, whereas no differences were detected between stress-sensitive and stress-resilient monkeys in the mRNA for the serotonin 1A autoreceptor or MAO-B. The determinants of stressfulness in women with SIA/FHA are incompletely understood, and this is certainly one of the variables that warrants further investigation.

Judd et al. reported that a single 2 mg dose of alprazolam, a GABA receptor agonist, decreased cortisol levels and increased LH pulse frequency from 0.8 to 2.0 pulses/8 h in women with stress-related anovulation, whereas its administration decreased LH pulse frequency in eumenorrheic women in the follicular phase. These data also suggest a possible indirect role for GABA neurons in the stress-induced neuromodulation of GnRH pulsatility. Not surprisingly, the neurochemistry of stress and SIA/FHA is far from simple, and firm mechanistic conclusions are not possible at present.

Behavioral, Nutritional, and Metabolic Influences on the Reproductive Axis

Identification of factors that activate the adrenal axis and suppress the thyroidal and ovarian axes can be challenging. It has been suggested that different stressors and behaviors activate somewhat different central mechanisms, such that the signals that alter hypothalamic function are specific to the type of stressor. Indeed, this variability may well explain in part why the neurochemistry of stress is so complex. Psychosocial dilemmas are seen as activating those central pathways subserving perception, whereas exercise and weight loss are generally viewed as disturbing metabolic regulation. Although it seems logical that specificity in the neural or peripheral cascades mediating responses to specific stressors exist, we currently have no method for clearly delineating psychogenic from metabolic stress. Psychogenic stress has a metabolic cost and metabolic stressors, such as food restriction and excessive exercise, are often initiated to cope with psychogenic stress. However, until proven otherwise, the safest assumption is that stress comes in flavors and that some individuals are more sensitive than others to the same stressor or set of stressors.

Behavioral Influences

A number of behavioral and other psychogenic factors including exercise, low weight and weight loss, affective and eating disorders, various personality characteristics, drug use, and external and intrapsychic stresses have been associated with functional hypothalamic hypogonadism. Given individual variation in metabolism, autonomic tone, habitus, aptitudes, attitudes, and psychological valences, what is stressful to one person may be more or less so to another. Therefore, it is not surprising to find behavioral heterogeneity in the pathogenesis of SIA/FHA. Any given stressor, when the dose is large enough, can probably activate the central neural pathways leading to the disruption of GnRH. In clinical research, the trend has been to study single stressors and to partition as separate populations women with exercise amenorrhea, anorexia nervosa, and idiopathic amenorrhea. Populations studied in clinical research settings may not be entirely representative of all women with SIA/FHA because research subjects must meet relatively strict inclusion and exclusion criteria. In general, women with SIA/FHA do not report or do not have an easily identified solitary stressor. Typically, there are multiple, seemingly minor, stressors, such as a combination of job or school pressures, poor eating habits, and increased energy expenditure through activity or exercise.

To understand the role of psychological variables, such as attitudes and expectations, in the pathogenesis of SIA/FHA, we compared three groups of women: those with eumenorrhea and demonstrable luteal adequacy; those with SIA/FHA unrelated to excessive exercise, weight loss, an eating disorder, drug use, or an affective disorder; and those with anovulation due to an identifiable organic cause. Being amenorrheic, regardless of cause, was associated with a compromised sense of psychological equilibrium as reported on psychometric inventories, but, as a group, only women with SIA/FHA differed from the other groups on scales that measured unrealistic expectations and dysfunctional attitudes (defined as attitudes likely to impair coping responses). For instance, women with SIA/FHA were both highly perfectionistic and sociotrophic (having a high need for social approval). Because perfectionism interferes with social approval or acceptance, one interpretation is that the concomitant high drive for perfectionism and sociotrophy

creates an intrapsychic conflict that women with SIA/FHA may not possess the appropriate coping skills to resolve. Another interpretation is that the expectation of simultaneously being perfect and garnering social approval is an unrealistic expectation of self and others. Our earlier study suggested that women with SIA/FHA had trouble relaxing and having fun, attributes that may further predispose them to value performance at the expense of psychological needs. Although women with SIA/FHA do not typically meet the criteria for an eating disorder, they display many attitudes and behaviors similar to women with eating disorders, such as a drive for thinness and disordered eating. What appears to separate women with undifferentiated SIA/FHA from those with an eating disorder is the degree of disturbance, including the degree of food restriction, weight loss, binging, and perfectionism. However, direct comparisons have not been conducted for any of these psychological attributes, so these interpretations are based on impressions from extant literature. The bottom line is that attitudes and expectations engender behaviors such as aberrant food intake and excessive exercise that further challenge the hypothalamus to maintain homeostasis. Whether the identifiable behavior is performance pressure, exercise, or irregular food intake, the end result is the same, that is, the disruption of the GnRH pulse generator to the extent that anovulation occurs. It stands to reason, then, that the key to recovery is to change both the behaviors and attitudes that have initiated and now sustain reduced GnRH drive. Once the hypothalamic GnRH pulse generator has been disrupted, it may take a prolonged duration of energy balance and improved psychological equilibrium for the hypothalamic allostasis to reverse and for ovulatory function to return.

The available data suggest that any behavior or expectation that concomitantly activates to a sufficient degree the HPA and HPT axes has the potential to disrupt the GnRH pulse generator. Thus, the list of behaviors and attitudes associated with the development of SIA/FHA is expected to be diverse and extensive. Generally, a mix of multiple, seemingly minor psychogenic and metabolic stressors appears to be more deleterious to reproductive function than a solitary stressor.

Nutritional and Metabolic Influences

Nutritional and metabolic signals play critical roles in the elaboration of homeostatic and allostatic responses to ongoing challenges and can influence the reproductive system at many levels. Research on overnutrition has focused primarily on specifying its effects on gonadal function, with some attention having been given to the secondary consequences of obesity on endometrial development and central hypothalamic drive. Few studies have characterized the role of nutritional signals in the modulation of reproductive function in men, but available evidence suggests that undernutrition is as deleterious to reproductive competency in men as it is in women.

Metabolic imbalance occurs when energy expenditure exceeds energy intake (negative energy balance). The brain is the most metabolically active tissue in the body, requiring 16 times more energy per unit mass than muscle tissue. Because humans have much bigger brains relative to body size than do other primates or other species, they use more of their daily energy intake than any other species to supply their brain; humans are estimated to use 25%, monkeys 8%, and rodents 5%. Indeed, as a species, humans may be uniquely sensitive to energy imbalance. Metabolism reflects the interactions of the neuroendocrine system (e.g., HPA and HPT axes), body composition (adipokines), and the enteric system (insulin, ghrelin, etc.). Given the multitude of redundant metabolic and nutritional signals communicating energy status to the brain, it is difficult to specify the independent role of any one factor or signal. Signals reflecting energy stores, recent nutritional information, and specific classes of nutrients are integrated in the central nervous system, particularly the hypothalamus, to coordinate energy intake and expenditure. Chronic energy deficiency alters thyroidal function to slow metabolism and correct negative energy balance. Food intake is influenced by availability, emotional state, social cues, and learned behaviors. Although peripheral signals convey information about energy stores and immediate energy availability to the hypothalamus, particularly the arcuate nucleus, these two categories of signals are not exclusive. For instance, leptin and insulin, which are actively transported across the blood–brain barrier, potentiate satiety signals. Thus, the level of any one signal *per se* may be less important than the action of that hormone, and the action is likely to be gated by the constellation of other metabolic signals. Putative orexigenic signals include ghrelin, NPY, orexins A and B, melanin-concentrating hormone (MCH), and agouti-related peptide. Putative anorectic and satiety signals include (but are not limited to) cortisol, CRH, insulin, glucose, resistin, leptin, proopiomelanocortin (POMC), cocaine- and amphetamine-regulated transcript (CART) peptide, peptide YY (PYY), and glucagon-like peptide (GLP)-1. Adipokines (adipocyte-derived hormones) implicated in energy regulation include leptin, adiponectin, and resistin. Leptin is the dominant long-term energy signal informing the brain of adipose

energy reserves; it also functions as a satiety signal. Adiponectin acts as an insulin-sensitizing agent by reducing hepatic glucose production and is reduced in obesity. Resistin is linked to insulin tolerance and decreases glucose uptake by adipocytes. Ghrelin, a 28-amino-acid acylated hormone that is also a growth hormone secretagogue, is produced by the gastrointestinal tract, especially the fundus of the stomach. Plasma ghrelin levels rise during fasting and immediately before anticipated mealtimes and then fall within an hour of food intake, suggesting that ghrelin is important for meal initiation. Given the plethora of nutritional and metabolic signals, specifying their independent effects is challenging, particularly in humans.

Relevant human studies on the role of these factors in reproduction are few. Depression, like stress, involves LHPA axis dysregulation and is associated with loss of appetite; resistin levels correlated with free cortisol levels and adiponectin correlated with insulin sensitivity in a study of depressed humans. In a rodent model, the central LHPA axis challenge of hypoxia and the LHPA axis feedback challenge of dexamethasone independently decreased body weight and increased adiponectin levels, whereas dexamethasone, but not hypoxia, increased resistin. In women, ghrelin levels were reported to be increased in anorexia nervosa and exercise amenorrhea. However, Troisi et al. observed lower ghrelin in women with anorexia or bulimia that correlated with serum cortisol levels. In these studies, subjects and controls differed markedly with regard to body weight and leptin levels, food intake was not assessed, and the diurnal pattern of ghrelin was not determined; for instance, De Souza et al. and Tanaka et al. obtained only a single morning's blood sample after an overnight fast and Misra et al. determined ghrelin levels only overnight. Of note, Miljec et al. reported ghrelin insensitivity in response to an infusion of ghrelin in women with anorexia nervosa compared to eumenorrheic women, but appetite was measured only by self-report and not by food intake. Women with SIA/FHA are more active than most eumenorrheic women and tend to consume less fat and more carbohydrates, but reproductive recovery from SIA/FHA did not require weight gain and was not associated with a change in standard metabolic variables such as leptin.

A balance between energy intake and expenditure is critical to survival. Every action, even thinking, has an energetic cost. Given the fundamental nature of metabolism, it is not surprising that organisms have evolved a complex and redundant signaling system to gate appetite, food-seeking behavior, and fuel storage. Although stress increases energetic demand, the acute behavioral response is often a blunting of appetite and a reduction in food intake so as to shift attention away from maintenance/sustenance needs and toward coping responses. Chronic reductions of appetite and energy intake in the face of an increase in energy demand probably represent appetite or metabolic allostasis; because of ensuing hypothalamic hypothyroidism, undernutrition may develop without weight loss. The exact mechanisms subserving the multiple associations between reproduction and nutrition and the applicability of animal studies to human syndromes remain to be determined. Given the considerations outlined here, including the high metabolic demands of the human brain, energetic imbalance may have more serious reproductive consequences for humans than for other animals.

Psychological states generate metabolic demands. Psychological states recognized as stressors activate the LHPA axis. When this activation is acute and limited, the adrenal releases a bolus of cortisol, a hormone with multiple metabolic properties. If the stressor is chronic and mild to moderate, then there is modest activation of the LHPA such that cortisol is mildly elevated but the circadian pattern of cortisol is preserved. With extreme unremitting stress, cortisol secretion is elevated and there is loss of the circadian pattern. Conditions that chronically activate the LHPA axis have been associated with a spectrum of reproductive impairment.

In summary, there is a plethora of metabolic signals that reflect the acute and chronic nutritional state of a given individual. These signals convey information to important tissues and organs and thereby engender homeostatic and allostatic responses. A key target tissue is the brain. Both undernutrition and overnutrition impact reproductive processes. Given the energetic demands of reproductive processes, however, undernutrition has the potential to induce more significant reproductive compromise.

Synergism among Stressors

The notion that stress comes in neurochemical flavors is supported by animal studies suggesting that there are subtle but distinct differences in the neuroendocrine responses to different stress paradigms. Further, neuroendocrine and metabolic responses to acute exercise were greater in men whose HPA axis did not suppress when they were given dexamethasone before the exercise challenge. These data indicate that the degree of HPA activation potentiates the neuroendocrine and metabolic responses to subsequent challenge. Conversely, Altemus et al. showed that lactating women are hyporesponsive to exercise challenge. Taken together, these data buttress the

notion that responses to a given stressor are gated not only by the stressor type but also by host factors, including hormonal and metabolic states. Thus, some individuals are clearly more reactive to similar stressors than others. Obviously, emotional valence and expectations also determine the extent to which psychosocial variables serve as a psychogenic stressors. Our preliminary investigations in women who developed SIA/FHA unrelated to weight loss, exercise, and definable psychiatric disorders indicated that a primary factor that distinguished women with SIA/FHA from those with definable causes of anovulation and those who were ovulatory was the presence of unrealistic expectations. Fioroni also found that women with SIA/FHA, compared to eumenorrheic women or those with other causes of anovulation, held more negative attributions about recent life events. Kirschbaum found that men who did not habituate when exposed to repeated psychogenic challenge viewed themselves as less attractive, had lower self-esteem, and reported being in a depressed mood more often. Apparently, unachievable ambitions or other cognitive distortions create vulnerability to life's inevitable challenges and probably heighten responsivity to metabolic stressors such as exercise or food restriction. Potentially, the converse is true. Energetic imbalance may augment reactivity to psychogenic stressors.

There can be no doubt that weight loss and exercise serve as metabolic stressors. In monkeys trained to run, it was shown that caloric supplementation reversed the anovulation induced by training. Interestingly, the monkeys did not spontaneously develop a compensatory increase in appetite and had to be bribed with colorful candy to consume more calories. On the other hand, modest dietary restriction accompanied by small amounts of exercise greatly increased the proportion of monkeys who become anovulatory when presented with social stress. A prospective study of unselected women demonstrated that exercise and weight loss caused anovulation. It is likely that sufficient exercise and weight loss, independent of psychogenic stress, can alter metabolism to the point that GnRH pulsatility is disrupted. Loucks and Thurma quantified the amount of energy restriction needed to impact GnRH drive by studying eumenorrheic women in the follicular phase. Energy balance was achieved by providing 45 kcal/kg of lean body mass (LBM)/day. Graded daily energy deficits of 10, 25, and 35 kcal/kg were then experimentally induced for 5 days. An energy deficit of 33% had no impact on LH pulse frequency, whereas an energy deficit of approximately 75% induced a decline in LH pulse frequency of approximately 40%. The induction of an energy deficit resulted in a graded increase in the 24-h cortisol level. At an energy availability of 10 kcal/kg of LBM (75% deficit), the mean 24-h cortisol was increased by approximately 30%, which is the amount of increase typically seen in women with SIA/FHA. Much like stress-sensitive monkeys, women whose luteal phase progesterone levels were lowest at the initiation of the energy restriction showed the greatest response to the imposed metabolic challenge. In most real-life situations, except for extreme circumstances such as war or famine, metabolic deficits are not imposed but, rather, initiated by individuals in response to self-imposed expectations including drive for thinness. Most women with SIA/FHA, when carefully evaluated, display more than one trait, state, or behavior capable of activating stress response cascades or inducing a mild metabolic deficit, but most do not have a profound metabolic deficit that alone would explain the reduction in central GnRH drive. Many of the behaviors, such as exercise, that independently suppress central reproductive drive but only at more extreme levels, may be initiated as coping responses to psychosocial dilemmas. Because of the synergism between metabolic and psychogenic stressors, a combination of multiple, small magnitude, mixed stressors may be potentially more disruptive of reproductive function than a single large stressor limited to one category.

To better understand how a combination of seemingly minor psychogenic and metabolic stressors might synergistically disrupt GnRH drive as demonstrated in our monkey model, we compared endocrine responses to submaximal exercise in women with SIA/FHA to those in eumenorrhea. Women with SIA/FHA displayed a larger increase in cortisol than ovulatory eumenorrheic women in response to exercise. Further, glucose responses between the two groups were divergent in that women with SIA/FHA showed a 10% decrease in glucose whereas eumenorrheic women had a modest 3% increase in glucose, presumably to cope with the energetic demand of exercise. Interestingly, these two groups did not differ at baseline with regard to cortisol or glucose levels. The decrement in glucose seen in SIA/FHA but not eumenorrheic women suggests latent metabolic compromise and indicates that SIA/FHA women are unable to meet energetic demands of ongoing activities. Further, it is likely that the drop in glucose activates the HPA axis and is at least partly responsible for the sustained hypercortisolemia characteristic of SIA/FHA. Because metabolic signals modulate GnRH pulsatility, exercise-induced metabolic imbalance also probably contributes to ongoing reproductive suppression. These results also reveal why the endocrine effects of a stressor, such as exercise, depend on the preexisting neuroendocrine and metabolic state of the individual.

Treatment Considerations

SIA may not be recognized unless the menstrual interval is markedly short, long, irregularly irregular, or absent. Likewise, luteal phase insufficiency due to decreased hypothalamic drive may not be noted unless infertility results. Even then, luteal insufficiency is notoriously difficult to document unless it is recurrent. Based on the foregoing concepts, we speculate that the more clinically evident the ovarian compromise, the greater is the hypothalamic challenge and the more profound are the associated adrenal and thyroidal derangements and sex steroid deprivation.

If a woman with functional hypothalamic hypogonadism is seeking to become pregnant, ovulation induction can be accomplished technically with the exogenous administration of pulsatile GnRH therapy or gonadotropins. The administration of exogenous GnRH therapy is advantageous insofar as it diminishes the risk of ovarian hyperstimulation and multiple gestation associated with gonadotropins. Clominphene citrate is an option, but it may be less effective because of its hypothalamic site of action and the hypothalamus of women with SIA/FHA is already not responding to decreased sex steroid secretion. Concerns have been raised that ovulation induction may place women with SIA/FHA at increased risk for premature labor and intrauterine growth restriction. The parenting skills of women with SIA/FHA may be impaired because they are already overwhelmed and stressed prior to pregnancy and delivery and their psychological precariousness may place their children at risk for poor psychosocial development. Further, a recent study showed that children born to mothers with clinically occult hypothyroidism due to autoimmune thyroiditis had a mean full-scale intelligence quotient that was seven points lower than the control population. The women with clinically silent hypothyroidism had a 30% reduction in thyroxine, which is roughly what is observed in women with SIA/FHA. Of critical importance is the fact that maternal thyroxine is the only source of fetal thyroxine in the first trimester and the predominant fetal source in the second and third trimesters. Because the fetal brain requires an appropriate amount of thyroxine for neurogenesis, even small deficits in thyroxine may induce neurodevelopmental deficits. Increased maternal cortisol may also have independent effects on fetal neurodevelopment and organogenesis. Recent evidence showed that severe stress, such as that associated with the unexpected death of a child, increased the risk of congenital anomalies of the cranial neural crest eightfold. Further, stress and its endocrine concomitants and even brief undernutrition have been implicated as a cause of preterm delivery. It is not known if the endocrine concomitants associated with SIA/FHA pose a similar risk, but this is clearly a potential hazard if ovulation induction is undertaken before amelioration of the allostatic changes in the adrenal and thyroidal axes.

A popular approach to a woman with SIA/FHA who is not seeking immediately to become pregnant is to offer her hormone replacement. This approach is based on the presumption that sex steroid deprivation is the primary therapeutic issue. There are inherent limitations with this approach, however. First, data indicate that exogenous sex steroid exposure does not fully promote bone accretion or cardioprotection in the presence of ongoing metabolic derangements. Second, the ongoing insults to the brain from chronic amplification of stress cascades go unchecked. Further, estrogen therapy does not correct hypothalamic hypothyroidism. In short, hormone therapy may mask potentially deleterious processes that are unlikely to be ameliorated by hormone exposure alone. It is critical to remember that SIA/FHA is more than a disorder of reduced GnRH secretion.

Hormone therapy *per se* is unlikely to be harmful, but more than hormone administration is needed – the stress process needs to be interrupted. Although psychopharmacological approaches have not been well studied, they probably could be used on an interim basis in special circumstances. The study by Judd suggested that a short course of alprazolam might be effective in reducing HPA activation and permitting hypothalamic-pituitary-ovarian (HPO) recovery. However, this approach is not recommended for a woman hoping to conceive. The optimal intervention is to reverse the stress process so that the hypothalamus recovers and gonadal function resumes. An integral goal of the treatment plan for women with functional hypothalamic hypogonadism is to help them identify and ameliorate the sources of psychogenic and metabolic stress and to provide emotional support while they learn coping mechanisms other than dieting or exercising. Nonpharmacological interventions such as stress management, relaxation training, or psychoeducation empower individuals by fostering self-care and competency. In this regard, nonpharmacological therapies have the potential to produce long-term benefits on psychological and, thereby, physical health. Behavioral therapies acknowledge the wisdom of the body and understand that SIA/FHA represents an endocrine adaptation that can be reversed with appropriate psychogenic and behavioral modifications.

Given these considerations, we recently studied whether CBT aimed at ameliorating problematic attitudes and behaviors facilitated ovarian recovery in

normal weight women with SIA/FHA. Women with SIA/FHA were randomly assigned to observation versus CBT groups. CBT consisted of 16 visits with a physician, therapist, or nutritionist during 20 weeks. The two groups were followed for return of menses for up to 8 weeks following the intervention. Regardless of menstrual pattern, estradiol and progesterone levels were monitored at weekly intervals for 4 weeks before and after observation or CBT. Approximately 88% of those who underwent CBT had evidence of ovulation, whereas only 25% of those in the observation group did. Figure 4 illustrates the recovery of sex steroid secretion in a woman in the CBT group who showed ovarian recovery and in a woman in the observation group who did not have recovery of ovarian function. Interestingly, when it occurred, ovarian recovery was not associated with significant weight gain. This does not mean that subjects did not alter food intake or energy expenditure, however. Improved nutrition generally restores the thyroidal

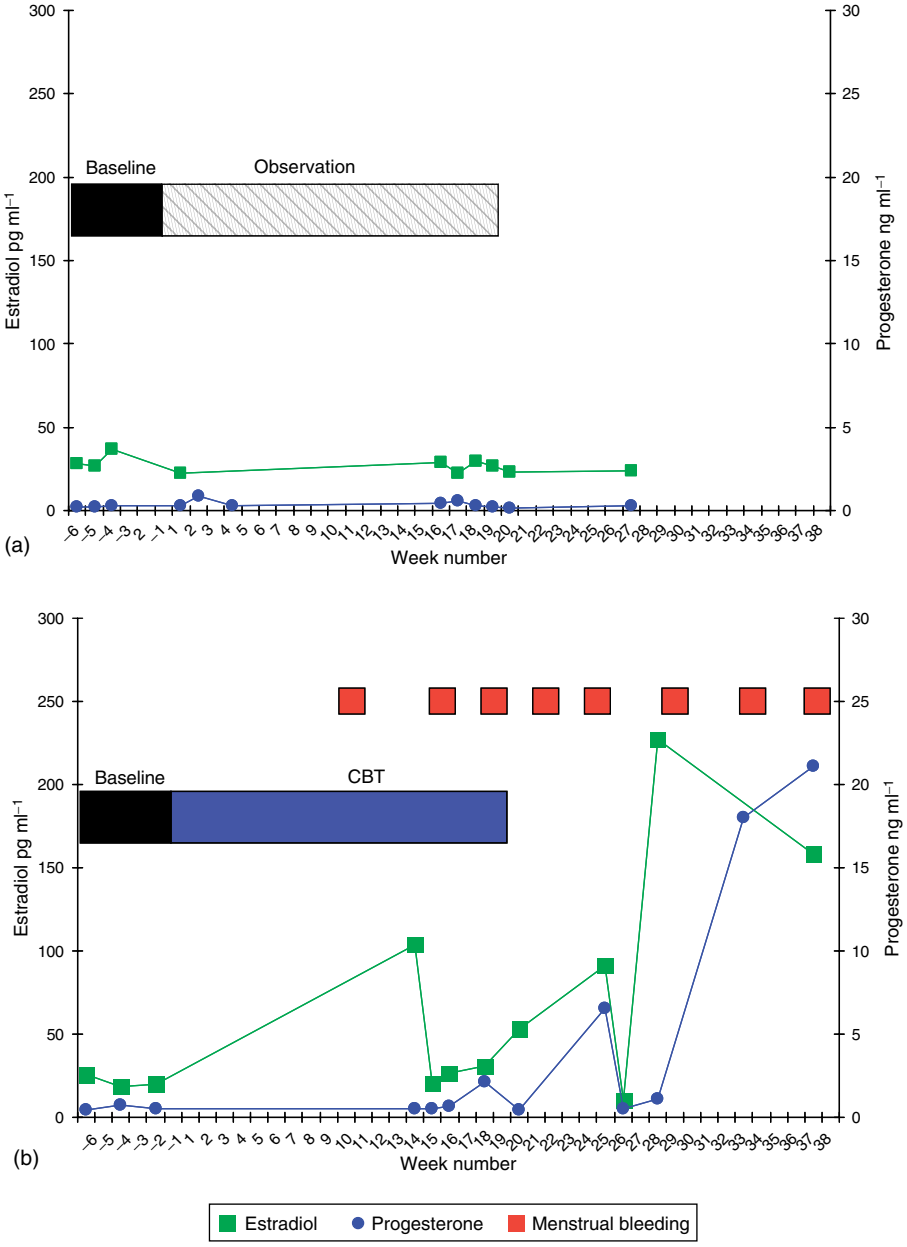


Figure 4 Serum estradiol and progesterone levels and menstrual bleeding in women with SIA/FHA. a, A patient who was observed and did not recover; b, a patient who was treated with cognitive-behavior therapy and recovered. CBT, cognitive-behavior therapy; SIA/FHA, stress-induced anovulation/functional hypothalamic amenorrhea. From Berga, S. L., Marcus, M. D., Loucks, T. L., et al. (2003), Recovery of ovarian activity in women with functional hypothalamic amenorrhea who were treated with cognitive behavior therapy, *Fertility and Sterility* **80**, 976–981.

axis, thereby leading to an increased basal metabolic rate and a need for more calories. Our earlier studies showed that women who experienced a spontaneous recovery from SIA/FHA were eucortisolemic, had increased but not fully restored LH pulsatility, and markedly increased TSH levels in the face of persistent reductions in T_3 and T_4 levels. We recently compared metabolic variables in women who did and who did not recover from SIA/FHA after CBT intervention and observed that ovarian recovery was associated with a decline in cortisol and increase in TSH but that thyroidal hormones and leptin levels were comparable between recovered and unrecovered women. Taken together, these findings suggest that the thyroidal axis, which gates metabolic tone, remains restrained by an as yet uncharacterized factor(s) after HPA and HPO recovery or recovers more slowly than the HPA and HPO axes, similar to the recovery timeline related to acute illness. Ultimately, behavioral and psychological interventions that address problematic behaviors and attitudes permit the resumption of ovarian function along with recovery of the LHPA and HPT axes. In short, full endocrine recovery offers better individual, maternal, and child health.

Summary

Stress is one of the most common and most commonly underappreciated causes of infertility and reproductive compromise in men and women. Stress-induced hypothalamic hypogonadism increases the acute and chronic health burden for individuals and their offspring. Our findings regarding the pathogenesis of SIA/FHA reveal a powerful synergism between metabolic imbalance and psychogenic challenge and underscore that seemingly mundane psychological issues evoke not only reproductive compromise but also concomitant neuroendocrine and metabolic allostasis. These mechanistic pathobiological insights have been harnessed to develop therapeutic options aimed at ameliorating the sustaining cognitive and behavioral variables. By showing that CBT reverses SIA/FHA, we can appreciate the inherent neural plasticity that forms the rationale for coupling mindfulness approaches with the judicious use of psychopharmacological and technical interventions.

Acknowledgments

Portions of the work presented herein were funded by grants from the National Institutes of Health (R01-MH50748 to SLB and RR-00046 to the General Clinical Research Center at the University of Pittsburgh).

See Also the Following Article

Allostasis and Allostatic Load.

Further Reading

Altemus, M., Deuster, P. A., Galliven, E., et al. (1985). Suppression of hypothalamic-pituitary-adrenal axis responses to stress in lactating women. *Journal of Clinical Endocrinology and Metabolism* **80**, 2954–2959.

Alvero, R., Kimzey, L., Sebring, N., et al. (1998). Effects of fasting on neuroendocrine function and follicle development in lean women. *Journal of Clinical Endocrinology and Metabolism* **83**, 76–80.

Ariyasu, H., Takaya, K., Tagami, T., et al. (2001). Stomach is major source of circulating ghrelin and feeding state determines plasma ghrelin-like immunoreactivity levels in humans. *Journal of Clinical Endocrinology and Metabolism* **86**, 4753–4758.

Badman, M. K. and Flier, J. S. (2005). The gut and energy balance: visceral allies in the obesity wars. *Science* **307**, 1909–1914.

Barr, C. S., Newman, T. K., Schwandt, M., et al. (2004). Sexual dichotomy of an interaction between early adversity and the serotonin transporter gene promoter variant in rhesus macaques. *Proceedings of the National Academy of Sciences USA* **101**, 12358–12363.

Barr, C. S., Newman, T. K., Shannon, C., et al. (2004). Rearing condition and rh5-HTTLPR interact to influence limbic-hypothalamic-pituitary-adrenal axis response to stress in infant macaques. *Biological Psychiatry* **55**, 733–738.

Becker, A. E., Grinspoon, S. K., Klibanski, A., et al. (1999). Eating disorders. *New England Journal of Medicine* **340**, 1092–1098.

Bennett, A. J., Lesch, K. P., Heils, A., et al. (2002). Early experience and serotonin transporter gene variation interact to influence primate CNS function. *Molecular Psychiatry* **7**, 118–122.

Berga, S. L. (2002). Disorders of gonadotropin secretion. In: Wass, J. & Shlet, S. M. (eds.) *Oxford textbook of e&ocrinology and diabetes*, pp. 1119–1129. Oxford University Press.

Berga, S. L. and Daniels, T. L. (1991). Use of the laboratory in disorders of reproductive neuroendocrinology. *Journal of Clinical Immunoassay* **14**, 23–28.

Berga, S. L., Daniels, T. L. and Giles, D. E. (1997). Women with functional hypothalamic amenorrhea but not other forms of anovulation display amplified cortisol concentrations. *Fertility and Sterility* **67**, 1024–1030.

Berga, S. L., Loucks, A. B., Rossmanith, W. G., et al. (1991). Acceleration of LH pulse frequency in functional hypothalamic amenorrhea by dopaminergic blockade. *Journal of Clinical Endocrinology and Metabolism* **72**, 151–156.

Berga, S. L., Loucks, T. L. and Cameron, J. L. (2001). Endocrine and chronobiological effects of fasting in women. *Fertility and Sterility* **75**, 926–932.

Berga, S. L., Loucks-Daniels, T. L., Adler, L. J., et al. (2000). Cerebrospinal fluid levels of corticotropin-releasing

hormone in women with functional hypothalamic amenorrhea. *American Journal of Obstetrics and Gynecology* **182**, 776–784.

Berga, S. L., Marcus, M. D., Loucks, T. L., et al. (2003). Recovery of ovarian activity in women with functional hypothalamic amenorrhea who were treated with cognitive behavior therapy. *Fertility and Sterility* **80**, 976–981.

Berga, S. L., Mortola, J. F., Girton, L., et al. (1989). Neuroendocrine aberrations in women with functional hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* **68**, 301–308.

Berga, S. L., Mortola, J. F. and Yen, S. S. C. (1988). Amplification of nocturnal melatonin secretion in women with functional hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* **66**, 242–244.

Berga, S. L. and Yen, S. S. C. (2004). Reproductive failure due to central nervous system-hypothalamic-pituitary dysfunction. In: Strauss, J. & Barbieri, R. (eds.) *Yen and Jaffe's reproductive endocrinology: physiology, pathophysiology, and clinical management* (5th edn., pp. 564–569). Philadelphia: Elsevier Saunders.

Bethea, C. L., Pau, F. K. Y., Fox, S., et al. (2005). Sensitivity to stress-induced reproductive dysfunction linked to activity of the serotonin system. *Fertility and Sterility* **83**, 148–155.

Bethea, C. L., Streicher, J. M., Mirkes, S. J., et al. (2005). Serotonin-related gene expression in female monkeys with individual sensitivity to stress. *Neuroscience* **132**, 151–166.

Biller, B. M. K., Federoff, J. H., Koenig, J. I., et al. (1990). Abnormal cortisol secretion and responses to corticotropin-releasing hormone in women with hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* **70**, 311–317.

Brundu, B., Loucks, T. L., Adler, L. J., et al. (2006). Increased cortisol in the cerebrospinal fluid of women with functional hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* **91**, 1561–1565.

Bullen, B. A., Skrinar, G. S., Betins, I. Z., et al. (1985). Induction of menstrual disorders by strenuous exercise in untrained women. *New England Journal of Medicine* **312**, 1349–1353.

Cameron, J. L., Nosbisch, C., Helmreich, D., et al. (1990). Reversal of exercise-induced amenorrhea in female cynomolgus monkeys (*Macasa fascicularis*) by increasing food intake. Paper presented at the 72nd annual meeting of the Endocrine Society, Atlanta, GA. Abstract # 1042.

Cameron, J. L., Weltzin, T. E., McConahan, C., et al. (1997). Slowing of pulsatile luteinizing hormone secretion in men after forty-eight hours of fasting. *Journal of Clinical Endocrinology and Metabolism* **73**, 35–41.

Caspi, A., Sudgen, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301**, 386–389.

Challis, J., Sloboda, D., Matthews, S., et al. (2000). Fetal hypothalamic-pituitary-adrenal (HPA) development and activation as a determinant of the timing of birth, and of postnatal disease. *Endocrine Research* **26**, 489–504.

Champoux, M., Bennett, A., Shannon, C., et al. (2002). Serotonin transporter gene polymorphism, differential early rearing, and behavior in rhesus monkey neonates. *Molecular Psychiatry* **7**, 1058–1063.

Cummings, D. E., Purnell, J. Q., Frayo, R. S., et al. (2001). A preprandiol rise in plasma ghrelin suggests a role in meal initiation in humans. *Diabetes* **50**, 1714–1719.

Daniels, T. L., Cameron, J. L., Marcus, M. D., et al. (1997). Recovery from functional hypothalamic amenorrhea (FHA) involves resolution of hypothalamic hypothyroidism. Paper presented at the 79th annual meeting of the Endocrine Society, Minneapolis, MN. Abstract #P1–378.

DeSouza, M. J., Miller, B. E., Loucks, A. B., et al. (1998). High frequency of luteal phase deficiency and anovulation in recreational women runners: blunted elevation in follicle-stimulating hormone observed during luteal-follicular transition. *Journal of Clinical Endocrinology and Metabolism* **83**, 4220–4232.

de Boo, H. A. and Harding, J. E. (2006). The developmental origins of adult disease (Barker) hypothesis. *Australian and New Zealand Journal of Obstetrics and Gyneacology* **46**, 4–14.

Drew, F. L. (1961). The epidemiology of secondary amenorrhea. *Journal of Chronic Diseases* **14**, 396–407.

Dudas, B. and Merchenthaler, I. (2006). Three-dimensional representation of the neurotransmitter systems of the human hypothalamus: inputs of the gonadotropin hormone-releasing hormone neuronal systems. *Journal of Neuroendocrinology* **18**, 79–95.

Fioroni, L., Fava, M., Genazzani, A. D., et al. (1994). Life events impact in patients with secondary amenorrhea. *Journal of Psychosomatic Research* **38**, 617–622.

Fries, H., Nilius, S. J. and Pettersson, F. (1974). Epidemiology of secondary amenorrhea. II: A retrospective evaluation of etiology with regard to psychogenic factors and weight loss. *American Journal of Obstetrics and Gynecology* **118**, 473–479.

Fuller, R. W. (1996). Serotonin receptors involved in regulation of pituitary-adrenocortical function in rats. *Behavioural Brain Research* **73**, 215–219.

Genazzani, A. D., Gamba, O. and Petraglia, F. (1998). Estrogen replacement therapy modulates spontaneous GH secretion but does not affect GHRH-induced GH response and low T3 syndrome in women with hypothalamic amenorrhea associated to weight loss. *Journal of Endocrinological Investigation* **21**, 353–357.

Giles, D. E. and Berga, S. L. (1993). Cognitive and psychiatric correlates of functional hypothalamic amenorrhea: a controlled comparison. *Fertility and Sterility* **60**, 486–492.

Golden, N. H., Lanzkowsky, L., Schebendach, J., et al. (2002). The effect of estrogen-progestin treatment on bone mineral density in anorexia nervosa. *Journal of Pediatric and Adolescent Gynecology* **15**, 135–143.

Goodman, R. L. (1989). Functional organization of the catecholaminergic neural system inhibiting luteinizing hormone secretion in anestrous ewes. *Neuroendocrinology* **47**, 203–250.

Grinspoon, S., Miller, K., Coyle, C., et al. (1999). Severity of osteopenia in estrogen-deficient women with anorexia nervosa and hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* **84**, 2049–2055.

Haddow, J. E., Palomaki, G. E., Allan, W. C., et al. (1999). Maternal thyroid deficiency during pregnancy and subsequent neuropsychological development of the child. *New England Journal of Medicine* **341**, 549–555.

Hansen, D., Lou, H. C. and Olsen, J. (2000). Serious life events and congenital malformations: a national study with complete follow-up. *Lancet* **356**, 875–880.

Heisler, L. E., Tumber, A. J., Reid, R. L., et al. (1994). Vasopressin mediates hypoglycemia-induced inhibition of luteinizing hormone secretion in the ovariectomized Rhesus monkey. *Neuroendocrinology* **60**, 297–304.

Holsboer, F. and Barden, N. (1996). Antidepressants and hypothalamic-pituitary-adrenocortical regulation. *Endocrine Review* **17**, 187–205.

Hurley, D. M., Brian, R., Outch, K., et al. (1984). Induction of ovulation and fertility in amenorrheic women by pulsatile low-dose gonadotropin-releasing hormone. *New England Journal of Medicine* **301**, 1069–1074.

Judd, S. J., Wong, J., Saloniklis, S., et al. (1995). The effect of alprazolam on serum cortisol and luteinizing hormone pulsatility in normal women and in women with stress-induced anovulation. *Journal of Clinical Endocrinology and Metabolism* **80**, 818–823.

Kaplan, J. R. and Manuck, S. B. (2004). Ovarian dysfunction, stress, and disease: a primate continuum. *ILAR Journal* **45**, 89–115.

Kasuya, E., Nyberg, C. L., Mogi, K., et al. (1999). A role of γ-aminobutyric acid (GABA) and glutamate in control of puberty in female rhesus monkeys: effect of an antisense oligodeoxynucleotide for GAD67 messenger ribonucleic acid and MK801 on luteinizing hormone-releasing hormone release. *Endocrinology* **140**, 705–712.

Khoury, S. A., Reame, N. E., Kelch, R. P., et al. (1987). Diurnal patterns of pulsatile luteinizing hormone secretion in hypothalamic amenorrhea: reproducibility and responses to opiate blockade and a α_2-adrenergic agonist. *Journal of Clinical Endocrinology and Metabolism* **64**, 755–762.

Kirschbaum, C., Prussner, J. C., Stone, A. A., et al. (1995). Persistent high cortisol responses to repeated psychological stress in a subpopulation of healthy men. *Psychosomatic Medicine* **57**, 468–474.

Kiss, J. and Halasz, B. (1985). Demonstration of serotonergic axons terminating on luteinizing hormone releasing hormone neurons in the preoptic area of the rat using a combination of immunocytochemistry and high resolution autoradiography. *Neuroscience* **14**, 69–78.

Klibanski, A. N., Biller, B. M. K., Schoenfeld, D. A., et al. (1994). The effects of estrogen administration on trabecular bone loss in young women with anorexia nervosa. *Journal of Clinical Endocrinology and Metabolism* **80**, 898–904.

Kojima, M., Hosoda, H., Date, Y., et al. (1999). Ghrelin is growth hormone releasing acylated peptide from stomach. *Nature* **402**, 656–660.

Kuljis, R. O. and Advis, J. P. (1989). Immunocytochemical and physiological evidence of a synapse between dopamine and luteinizing hormone releasing hormone-containing neurons in the ewe median eminence. *Endocrinology* **124**, 1579–1581.

Leonard, W. R. and Robertson, M. L. (1994). Evolutionary perspectives on human nutrition: the influence of brain and body size on diet and metabolism. *American Journal of Human Biology* **6**, 77–88.

Leranth, C., MacLusky, N. J., Shanabrough, M., et al. (1988). Catecholaminergic innervation of luteinizing hormone-releasing hormone and glutamic acid decarboxylase immunopositive neurons in the rat medial preoptic area. *Neuroendocrinology* **48**, 581–602.

Lesch, K. P., Bengel, D., Heils, A., et al. (1996). Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* **274**, 1527–1531.

Li, S. and Pelletier, G. (1996). Further studies on the mechanism of action of the endogenous benzodiazepine receptor ligand octadecaneuropeptide on gonadotropin-releasing hormone gene expression in the rat brain. *Neuroendocrinology* **64**, 79–84.

Lopez, J. F., Chalmers, D. T., Little, K. Y., et al. (1998). Regulation of serotonin, glucocorticoid and mineralocorticoid receptor in rat and human hippocampus: implications for the neurobiology of depression. *Biological Psychiatry* **43**, 547–573.

Loucks, A. B., Laughlin, G. A., Mortola, J. F., et al. (1992). Hypothalamic-pituitary-thyroidal function in eumenorrheic and amenorrheic athletes. *Journal of Clinical Endocrinology and Metabolism* **75**, 514–518.

Loucks, A. B., Mortola, J. F., Girton, L., et al. (1989). Alternations in the hypothalamic-pituitary-ovarian and hypothalamic-pituitary-adrenal axes in athletic women. *Journal of Clinical Endocrinology and Metabolism* **68**, 402–411.

Loucks, A. B. and Thurma, J. R. (2003). Luteinizing hormone pulsatility is disrupted at a threshold of energy availability in regularly menstruating women. *Journal of Clinical Endocrinology and Metabolism* **88**, 297–311.

Loucks, T. L., Dube, J., Laychak, K., et al. (2004). Metabolic and endocrine responses to submaximal exercise challenge in women with functional hypothalamic amenorrhea. Paper presented at the 51st annual meeting of the Society of Gynecologic Investigation, Houston, TX. Abstract #605.

Lupien, S. J., Nair, N. P., Briere, S., et al. (1999). Increased cortisol levels and impaired cognition in human aging: implication for depression and dementia in later life. *Reviews of Neuroscience* **10**, 117–139.

MacLusky, N. J., Naftolin, F. and Leranth, C. (1988). Immunocytochemical evidence for direct synaptic connections between corticotropin-releasing factor (CRF) and gonadotropin-releasing hormone (GnRH)-containing neurons in the preoptic area of the rat. *Brain Research* **439**, 381–395.

Mancini, F., Basanta-Henry, P., Corominal, A. G., et al. (2006). Do metabolic variables gate reproductive recovery in functional hypothalamic amenorrhea? Paper presented at the 88th annual meeting of the Endocrine Society, Boston, MA. Abstract # P2–531.

Marcus, M. D., Loucks, T. L. and Berga, S. L. (2001). Psychological correlates of functional hypothalamic amenorrhea. *Fertility and Sterility* **76**, 310–316.

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine* **338**, 171–179.

McEwen, B. S. (2002). Sex, stress, and the hippocampus: allostasis, allostatic load, and the aging process. *Neurobiology of Aging* **23**, 921–939.

McMillen, I. C., Schwartz, J., Coulter, C. L., et al. (2004). Early embryonic environment, the fetal pituitary-adrenal axis, and the timing of parturition. *Endocrine Review* **30**, 845–850.

Mikkelsen, J. D., Hay-Schmidt, A. and Kiss, A. (2004). Serotonergic stimulation of the rat hypothalamo-pituitary-adrenal axis: interaction between 5-HT1A and 5-HT2A receptors. *Annals of the New York Academy of Sciences* **1018**, 65–70.

Miljic, D., Pekic, S., Djurovic, M., et al. (2006). Ghrelin has partial or no effect on appetite, growth hormone, prolactin, and cortisol release in patients with anorexia nervosa. *Journal of Clinical Endocrinology and Metabolism* **91**, 1491–1495.

Miller, D. S., Reid, R. R., Cetel, N. S., et al. (1983). Pulsatile administration of low-dose gonadotropin-releasing hormone: ovulation and pregnancy in women with hypothalamic amenorrhea. *Journal of the American Medical Association* **250**, 2937–2941.

Misra, M., Miller, K. K., Kuo, K., et al. (2005). Secretory dynamics of ghrelin in adolescent girls with anorexia nervosa and healthy adolescents. *American Journal of Physiology, Endocrinology and Metabolism* **289**, E347–E356.

Morreale de Escobar, G., Obregon, M. J. and Escobar del Rey, F. (2000). Is neuropsychological development related to maternal hypothyroidism or to maternal hypothyroxinemia? *Journal of Clinical Endocrinology and Metabolism* **85**, 3975–3987.

Munoz, M. T., Morande, G., Garcia-Centenera, J. A., et al. (2002). The effects of estrogen administration on bone mineral density in adolescents with anorexia nervosa. *European Journal of Endocrinology* **146**, 45–50.

Murphy, D. L., Li, Q., Engel, S., et al. (2001). Genetic perspectives on the serotonin transporter. *Brain Research Bulletin* **56**, 487–494.

Nachtigall, L. B., Boepple, P. A., Pralong, F. P., et al. (1997). Adult-onset idiopathic hypogonatropic hypogonadism – a treatable form of male infertility. *New England Journal of Medicine* **336**, 410–415.

Petraglia, F., Vale, W. and Rivier, C. (1986). Opioids act centrally to modulate stress-induced decrease in luteinizing hormone in the rat. *Endocrinology* **119**, 2445–2450.

Petrides, J. S., Mueller, G. P., Kalogeras, K. T., et al. (1994). Exercise-induced activation of the hypothalamic-pituitary-adrenal axis: marked differences in the sensitivity to glucocorticoid suppression. *Journal of Clinical Endocrinology and Metabolism* **79**, 377–383.

Pike, K. M. and Rodin, J. (1991). Mothers, daughters, and disordered eating. *Journal of Abnormal Psychology* **100**, 198–204.

Prevot, V., Croix, D., Bouret, S., et al. (1999). Definitive evidence for the existence of morphological plasticity in the external zone of the median eminence during the rat estrous cycle: implication of neuro-glio-endothelial interactions in gonadotropin-releasing hormone release. *Neuroscience* **94**, 809–819.

Quigley, M. E., Sheehan, K. L., Casper, R. F., et al. (1980). Evidence for increased dopaminergic and opioid activity in patients with hypothalamic hypogonadotropic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* **50**, 949–954.

Raff, H. and Bruder, E. D. (2006). Adiponectin and resistin in the neonatal rat: effects of dexamethasone and hypoxia. *Endocrine* **29**, 341–344.

Reame, N. E., Sauder, S. E., Case, G. D., et al. (1985). Pulsatile gonadotropin secretion in women with hypothalamic amenorrhea: evidence that reduced frequency of gonadotropin-releasing hormone secretion is the mechanism of persistent anovulation. *Journal of Clinical Endocrinology and Metabolism* **61**, 851–858.

Reindollar, R. H., Novak, M., Tho, S. P. T., et al. (1986). Adult-onset amenorrhea: a study of 262 patients. *American Journal of Obstetrics and Gynecology* **155**, 531–543.

Romero, L. M., Plotsky, P. M. and Sapolsky, R. M. (1993). Patterns of adrenocorticotropin secretory release with hypoglycemia, novelty, and restraint after colchicine blockade of axonal transport. *Endocrinology* **132**, 199–204.

Sapolsky, R. M. (2000). Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. *Archives of General Psychiatry* **57**, 925–935.

Sapolsky, R. M. (2005). The influence of social hierarchy on primate health. *Science* **308**, 648–652.

Sapolsky, R. M. and Krey, L. C. (1988). Stress-induced suppression of luteinizing hormone concentrations in wild baboons: role of opiates. *Journal of Clinical Endocrinology and Metabolism* **66**, 772–776.

Schwartz, M. W. and Morton, G. J. (2002). Keeping hunger at bay. *Nature* **418**, 595–597.

Shanan, J., Brezinski, A., Sulman, F., et al. (1965). Active coping behavior, anxiety, and cortisol steroid excretion in the prediction of transient amenorrhea. *Behavioral Sciences* **10**, 461–465.

Spratt, D. I., Longcope, C., Cox, P. M., et al. (1993). Differential changes in serum concentrations of androgens and estrogens (in relation with cortisol) in postmenopausal women with acute illness. *Journal of Clinical Endocrinology and Metabolism* **76**, 1542–1547.

Stein-Behrens, B., Mattson, M. P., Chang, I., et al. (1994). Stress exacerbates neuron loss and cytoskeletal pathology in the hippocampus. *Journal of Neuroscience* **14**, 5373–5380.

Suh, B. Y., Liu, J. H., Berga, S. L., et al. (1988). Hypercortisolism in patients with functional hypothalamic amenorrhea. *Journal of Clinical Endocrinology and Metabolism* **66**, 733–739.

Sullivan, S. D., Howard, L. C., Clayton, A. H., et al. (2002). Serotonergic activation rescues reproductive function in fasted mice: does serotonin mediate the metabolic effects of leptin on reproduction? *Biology of Reproduction* **66**, 1702–1706.

Tanaka, M., Naruo, T., Yasuhara, D., et al. (2003). Fasting plasma ghrelin levels in subtypes of anorexia nervosa. *Psychoneuroendocrinology* **28**, 829–835.

Terasawa, E. (1995). Control of luteinizing-hormone releasing hormone pulse generation in nonhuman primates. *Cellular and Molecular Neurobiology* **15**, 141–164.

Thind, K. K., Boggan, J. E. and Goldsmith, P. C. (1991). Interactions between vasopressin- and gonadotropin-releasing-hormone-containing neuroendocrine neurons in the monkey supraoptic nucleus. *Neuroendocrinology* **53**, 287–297.

Thind, K. K. and Goldsmith, P. C. (1989). Corticotropin-releasing factor neurons innervate dopamine neurons in the periventricular hypothalamus of juvenile macaques. *Neuroendocrinology* **50**, 351–358.

Thind, K. K. and Goldsmith, P. C. (1998). Infundibular gonadotropin-releasing hormone neurons are inhibited by direct opioid and autoregulatory synapses in juvenile monkeys. *Neuroendocrinology* **47**, 203–216.

Troisi, A., DiLorenzo, G., Lega, I., et al. (2005). Plasma ghrelin in anorexia, bulimia, and binge-eating disorder; relations with eating pattern and circulating concentrations of cortisol and thyroid hormones. *Neuroendocrinology* **81**, 259–266.

Van der Spuy, Z. M., Steer, P. J., McCusker, M., et al. (1988). Outcome of pregnancy in underweight women after spontaneous and induced ovulation. *British Medical Journal* **296**, 962–965.

Vigersky, R. A., Anderson, A. E., Thompson, R. H., et al. (1977). Hypothalamic dysfunction in secondary amenorrhea associated with simple weight loss. *New England Journal of Medicine* **297**, 1141–1145.

Vitale, M. L. and Chiocchio, S. R. (1993). Serotonin, a neurotransmitter involved in regulation of luteinizing hormone release. *Endocrine Review* **14**, 480–493.

Wadhwa, P. D., Sandman, C. A. and Garite, T. J. (2001). The neurobiology of stress in human pregnancy: implications for prematurity and development of the fetal nervous system. *Progress in Brain Research* **133**, 131–142.

Warren, M. P. and Fried, J. L. (2001). Hypothalamic amenorrhea: the effects of environmental stresses on the reproductive system, a central effect of the central nervous system. *Endocrinology and Metabolism Clinics of North America* **30**, 611–629.

Warren, M. P., Miller, K. K., Olson, W. H., et al. (2005). Effects of an oral contraceptive (norgestimate/ethinyl estradiol) on bone mineral density in women with hypothalamic amenorrhea and osteopenia: an open-label extension of a double-blind, placebo-controlled study. *Contraception* **72**, 206–211.

Weber-Hamann, B., Kratzsch, J., Kopf, D., et al. (2006). Resistin and adiponectin in major depression: the association with free cortisol and effects of antidepressant treatment. *Journal of Psychiatric Research,* e-pub. ahead of print.

Whitnall, M. H. (1989). Stress selectively activates the vasopressin-containing subset of corticotropin-releasing hormone neurons. *Neuroendocrinology* **50**, 702–707.

Williams, C. L., Nishihara, M., Thalabard, J. C., et al. (1990). Corticotropin-releasing factor and gonadotropin-releasing hormone pulse generator activity in the rhesus monkey. *Neuroendocrinology* **52**, 133–137.

Williams, N. L., Berga, S. L. and Cameron, J. L. (1997). Mild metabolic stress potentiates the suppressive effects of psychological stress on reproductive function in female cynomolgus monkeys. Paper presented at the 79th annual meeting of the Endocrine Society, Minneapolis, MN. Abstract #P1–367.

Stress Management and Cardiovascular Disease

D Lane
Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK
D Carroll
University of Birmingham, Birmingham, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D Lane and D Carroll, volume 3, pp 527–531, © 2000, Elsevier Inc.

Stress Management Training
Essential Hypertension
Angina Pectoris
Myocardial Infarction
Conclusions

Glossary

Angina pectoris	Chest pains, often radiating to the arms, shoulders, and jaw, arising from reduced blood flow to the heart as a result of coronary atherosclerosis.
Angiography	A technique for measuring the extent of coronary atherosclerosis.
Atherosclerosis	An accumulation of fatty deposits causing hardening and narrowing of the arteries.
Cardiac rehabilitation	A combination of lifestyle education, exercise, and, increasingly, stress management training directed at patients suffering from coronary heart disease. The aim of cardiac rehabilitation is to facilitate physical, psychological, social, and emotional recovery.

Essential hypertension	Consistently high blood pressure without any apparent organic or medical cause.
Meta-analysis	Statistical techniques for combining and analyzing the outcomes of several studies.
Myocardial infarction (MI)	Heart attack.
Stress management training (SMT)	A range of techniques with the overall aim of training people to recognize and cope with psychological stress.
Thrombosis	Formation of a blood clot.

Stress Management Training

While stress management training (SMT) denotes a wide variety of practices concerned with improving individuals' recognition of and ability to cope with stress, most of the applications in cardiovascular disease include at least some of the following components: education about the nature of stress and its role in cardiovascular disease; identification of behavioral risk factors and counseling for lifestyle change; relaxation and its application during stress exposure; cognitive restructuring, in which positive and adaptive thoughts and beliefs are substituted for negative and maladaptive ones; and anger and hostility control when thought germane. Generally, SMT programs are undertaken with small groups of patients to maximize efficiency, trade on group motivational processes, and help generate social support.

Essential Hypertension

Essential hypertension is the presence of high blood pressure when there is no detectable medical or organic cause. The vast majority of instances of high blood pressure are of this sort. Hypertension is currently defined as a blood pressure of 140/90 mmHg or greater. Essential hypertension is a substantial problem. It is estimated that as much as 15% of the adult populations of Western countries exhibits high blood pressure, and various epidemiological studies testify that as blood pressure increases, life expectancy decreases, largely from an increased likelihood of stroke and myocardial infarction (MI). Given that psychological stress has been implicated in hypertension and that antihypertensive drugs often have undesirable side effects, SMT emerges as a potential adjunctive or even alternative therapy.

There have been more than 50 publications since the mid-1970s reporting attempts to reduce high blood pressure using some variant of SMT. Considered together, they paint a less than consistent picture. This is perhaps not surprising. First, although some form of relaxation training characterizes most interventions in this context, other components of SMT are frequently missing. While generalizations are difficult, interventions that have included more SMT components have tended to prove more effective. Second, studies differ in design. Although many have random allocation to SMT and control, control conditions vary between no concerted treatment and carefully engineered attention placebo. SMT appears to fare better compared to minimal interventions than when compared to placebo or active alternative treatments. A meta-analysis by Eisenberg and colleagues concluded than SMT was "superior to no therapy but not superior to credible sham techniques or to self-monitoring alone." (Eisenberg et al., 1993: 964). Third, studies differ in terms of patients' medication status. Results are difficult to interpret when patients are taking antihypertensive medication, since reductions in blood pressure associated with SMT could primarily reflect increased adherence to antihypertensive medication regimens. That may not be a bad thing. Without intervention, hypertensives often show low adherence. Fourth, studies vary in terms of whether they have included follow-up blood pressure assessment, although those that have generally find that gains associated with SMT programs are, to a reasonable extent, maintained at follow-up. Finally, most studies have restricted their outcome measures to resting blood pressure; few have included ambulatory blood pressure assessments during everyday life. Those that have do not uniformly find an advantage for SMT over control treatments.

A serious methodological weakness in virtually all studies to date is the failure to establish a stable baseline blood pressure, by repeated blood pressure determinations, against which to compare the effects of intervention. The phenomenon of white coat hypertension serves to elevate some individual's blood pressure, particularly in clinical settings. The implication of this phenomenon and the failure to establish stable baselines are clear; they are likely to flatter the potential hypotensive effects of interventions. An exception to the usual trend of summary baseline assessment is the study by Johnston and co-workers. Patients were referred to the study by general practitioners if they registered diastolic blood pressure between 90 and 104 mmHg on at least two out of three occasions of measurement. None were taking antihypertensive medication. Prior to allocation to a SMT treatment or an active placebo condition, patients, following training, took their own blood pressures twice daily for 3 months. During this 3-month baseline, blood pressure dropped from referral to start of treatment from 152/98 mmHg to 140/93 mmHg for the SMT group and from 152/98 mmHg to

140/92 mmHg for the control group. Neither SMT nor the control treatment affected any further reductions in blood pressure. SMT and control patients did not differ in follow-up blood pressure or in blood pressure response to a stressful interview. A major implication of this carefully conducted study is that previously reported blood pressure reductions associated with SMT may have reflected adaptation effects. SMT might merely help the process of adaptation to measurement, lowering the blood pressure readings of individuals who are not well adapted to blood pressure measurement procedures. This would render SMT of doubtful clinical significance. Given the importance of the Johnston and colleagues' study, replication is critical. However, it is critical that future studies include prolonged baseline assessment to counter the potentially confounding effects of adaptation to measurement.

Angina Pectoris

Coronary heart disease (CHD) remains the most common cause of death in Western countries for both men and women. The failure of traditional risk factors to provide a complete account of the etiology and course of coronary heart disease has fuelled speculation that other, possibly psychological, factors are involved. Key among these is psychological stress. A frequent expression of coronary heart disease is angina pectoris, experienced as a severe pain in the chest often radiating to the arms, shoulders, and jaw. Angina is a reflection of atherosclerosis of the coronary arteries limiting the supply of circulatory oxygen to the myocardium and, in consequence, reducing the capacity of the heart to deal with increases in load, as, for example, during exercise and psychological stress.

It is therefore surprising that few studies have evaluated the effects of stress management on angina. A notable exception is the Lifestyle Heart Trial. Angina patients were randomly assigned to a lifestyle change intervention or a usual care control group. Among other things, the lifestyle change group patients undertook a fairly comprehensive SMT program, practicing for at least 1 h a day, using the hour-long audiocassette tape provided, for a year. Their data indicate very good compliance. The major outcome measure was the status of patients' coronary arteries as determined by angiography. While there was a worsening of coronary atherosclerosis in the usual care control group over the course of the year, there was evidence of regression of atherosclerosis in the lifestyle change group. However, it is difficult to attribute these impressive gains to SMT, since patients in the lifestyle change condition also switched to a low-fat vegetarian diet and undertook regular vigorous exercise.

Some evidence that SMT may be a key player in these effects comes from the work of Bundy and associates. In the first of two studies, angina patients were assigned either to a comprehensive SMT program lasting 7 weeks or to routine care. Compared to patients receiving routine care, those undergoing SMT reported briefer angina attacks and reduced reliance on medication and showed increased tolerance of exercise, that is, they were able to exercise for a longer period of time before ischemia or pain curtailed their exertions. In a second study, the effects of SMT *per se* were less impressive. Nevertheless, patients who underwent SMT in combination with light exercise faired particularly well. Compared with patients receiving only SMT or usual care, they showed greater exercise tolerance, and compared with patients receiving only exercise or usual care, they registered less frequent angina attacks and were less reliant on medication.

More recently, Lewin and colleagues evaluated the effects of a cognitive-behavioral management program, called the Angina Plan, on psychological adjustment in patients with newly diagnosed angina pectoris. Patients were randomized to receive either the self-help manual (Angina Plan) together with an audio-taped relaxation program or a nurse-led education session, identifying patients' risk factors for CHD, with additional written information. At the 6-month follow-up, patients receiving the Angina Plan demonstrated significantly greater reductions in anxiety and depression, diminished frequency of angina, and fewer physical limitations than patients receiving only the educational session. Although the Angina Plan encompasses SMT techniques, it also consists of a combination of exercise and lifestyle changes. Therefore, it is difficult to attribute the changes demonstrated solely to SMT. However, considered overall, the available data would seem to invite further study of the potential of SMT in controlling angina pectoris.

Myocardial Infarction

Heart attack, or MI, is the most common traumatic consequence of coronary artery disease. It results from blockages in the coronary arteries sufficiently severe to deprive the myocardium of oxygen and damage the muscle to an extent that it ceases to function. MI reflects both atherosclerosis, in which the coronary arteries become narrowed by the formation of fatty deposits or plaques, and thrombosis, the formation of blood clots in the arteries, which serve to fully occlude the already compromised blood flow. However, substantial numbers of patients survive heart attacks, and the widespread administration of

anticoagulant drugs in recent years has increased survival rates substantially. As a consequence, effective cardiac rehabilitation has become an increasingly important health services objective. The overall level of provision has increased markedly in the last 10 years in Western countries. For example, in the UK, surveys reveal that there are now some 300 cardiac rehabilitation programs and that around 30% of major hospitals offer programs. While cardiac rehabilitation was traditionally restricted to basic health education and supervised regular exercise, some form of SMT is increasingly part of the package.

There are sound prime facie reasons for including SMT in this context. First, a heart attack is an exceedingly distressing experience for most patients, and levels of depression and anxiety following MI are frequently observed. Second, depressed and anxious cardiac patients are less likely to undertake lifestyle changes, adhere to medication, and exercise, and they also experience greater difficulties in coping with stress. Third, recent evidence indicates that post-MI levels of distress predict subsequent cardiac morbidity and mortality in some, but not all, studies. Fourth, psychological distress appears to be the most important predictor of hospitalization costs following a cardiac event, with distressed cardiac patients accruing four times the cost of nondistressed patients for nonpsychiatric medical interventions.

Early systematic evaluations of cardiac rehabilitation were restricted to programs that were largely exercise based and did not include SMT. Meta-analyses suggest that while such programs had no effects on the recurrence of nonfatal MIs, they were associated with reductions in mortality of 20 to 25%. However, since the studies included in these meta-analyses were generally undertaken with patients who were at low risk for recurrence of MI, the scope for demonstrating benefit was perhaps limited.

Most cardiac rehabilitation programs have included some form of SMT, although, as we shall see, in some instances the provision contains few, if any, of the customary elements of SMT. These have also been subjected to meta-analytic review. Linden and colleagues identified 23 randomized control trials of cardiac rehabilitation programs that contained at least some of the constituents of SMT. In total, these studies afforded data on 2024 patients who had received SMT and 1156 control patients who, for the most part, had received only usual care. The patients given SMT exhibited more substantial reductions in psychological distress, systolic blood pressure, and cholesterol than the controls. Critically, relative to the control condition, SMT was associated with a 40% greater chance of survival during the first 2 years following MI, although not subsequently. The risk of recurrence of MI was also associated with a substantially reduced risk of MI recurrence.

Since this meta-analysis, four randomized control trials with substantial patient numbers have been published. All four were concerned with alleviating patients' distress following MI. In the Montreal Heart Attack Readjustment Trial (M-Hart), patients' distress was monitored by monthly telephone calls for a year. Those judged to be distressed received a home visit from a nurse who offered emotional support, education, or referral to a general practitioner or cardiologist, as deemed appropriate. Patients in the program were compared to those receiving usual care. The intervention had only minimal effect on distress and no impact on overall survival at 1 year. Indeed, for female patients, cardiac and all-cause mortality tended to be higher in the intervention group. A similar sort of intervention strategy was used by Barr-Taylor and associates and again was compared to usual care. Nurses assessed patients' distress by telephone, and those registering high levels of distress were referred to mental health specialists, although no data were available on how many patients were referred and the nature of any treatment they received. While patients' distress lessened over the 1-year follow-up, there were no group differences. There were insufficient deaths to permit statistical comparison, although the deaths that occurred were split fairly evenly between groups. These two studies have been subject to a number of criticisms. Most importantly, the interventions contained little systematic SMT and, indeed, are barely recognizable as modern cardiac rehabilitation programs. It is likely that they were simply too minimal to influence cardiac health status. Predicated as they were on the reasonable presumption that alleviating patients' distress would yield dividends for recurrence and mortality, the failure of these studies to differentially ameliorate distress makes the absence of significant cardiac health effects unsurprising.

The third SMT intervention also failed to counter patients' distress following MI, again making the absence of effects on cardiac morbidity and mortality entirely predictable. However, the failure of the intervention to alleviate distress in this instance is more difficult to understand. Jones and West randomly allocated 2328 MI patients either to a reasonably comprehensive SMT program or to routine care. The SMT program comprised anxiety reduction techniques, relaxation training, instruction in coping skills, and attempts to promote a more positive adjustment to illness. As indicated previously, there were no differences between groups at 1-year follow-up either in reported distress or in clinical complications and mortality. There are two possible explanations for these

negative findings. First, many patients in the usual care group had good outcomes, which would serve to minimize any advantage for SMT. Second, it is clear that the intervention was not sufficiently powerful to affect psychological change in the most distressed patients. This study suggests, however, that routine psychological screening and intervention may be a useful first step but that more intensive SMT would be required for patients exhibiting extreme emotional distress.

The final randomized controlled trial, Enhancing Recovery in Coronary Heart Disease (ENRICHD), enrolled 2481 patients who had major or minor depression after their MI to receive either cognitive-behavioral therapy (CBT; n = 1238), with the option of adjunctive antidepressant medication (n = 249), or usual care (n = 1243). The main outcome measures were the assessment of depressive symptoms at 6 months, recurrent MI, and all-cause mortality. The intervention consisted of 6 months or less of both individual and group CBT, with variable numbers of sessions per patient. The intervention reduced depression at 6 months but had no significant effect on re-infarction or all-cause mortality, assessed after a mean follow-up of 41 months.

A recent Cochrane review examined the impact of nonpharmacological psychological interventions for adults with CHD (MI, coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty, angina, and angiographically defined CHD) on all-cause and CHD-related mortality, recurrent cardiac events, and symptoms of anxiety, depression, and stress, compared with adults receiving usual care or no intervention. Thirty-six trials were identified, 18 specifically examining SMT interventions (defined as the use of specific cognitive-behavioral strategies used to help the patient reduce or manage their stress). Of these 18 SMT trials, 11 reported on SMT interventions plus other rehabilitation interventions and seven reported on SMT interventions alone.

Rees and colleagues' systematic review concluded that there was no evidence of any benefit from SMT interventions in reducing all-cause or cardiac mortality. Further, the effects on nonfatal re-infarction were inconclusive. This conclusion was mainly influenced by two large negative randomized controlled trials, one of SMT and the other of CBT. However, trials with a SMT component did demonstrate a reduction in anxiety and depression following intervention. Nevertheless, caution is warranted in interpreting this result, given that it is based on small trials that included a combination of interventions. Accordingly, it is difficult to attribute with certainty such anxiolytic and antidepressive effects to SMT.

Conclusions

Stress management training has now been applied to a number of cardiovascular problems. The largest amount of research attention has been directed toward patients following MI. The data are extremely mixed, and two large trials have yielded disappointing results in terms of cardiological outcomes. Nevertheless, it appears that mixed programs of SMT and exercise rehabilitation produce psychological benefits comparable to those reported for several pharmacological interventions. Further research is required to identify the most effective intervention strategies. It is likely that the greatest dividends will accrue from individually tailored interventions, directing SMT to those patients most in need. Although angina has attracted less study, early indications are that SMT, particularly when combined with exercise, can produce improvements in symptoms and reduce reliance on medication and may even stem atherosclerotic deterioration of the coronary arteries. Finally, the indications for SMT in hypertension are far from compelling. It is possible that SMT acts simply to increase the rate of adaptation to repeated blood pressure measurement.

See Also the Following Articles

Cardiovascular System and Stress; Cognitive Behavioral Therapy; Depression and Coronary Heart Disease; Heart Disease/Attack; Cardiovasular Disease, Stress and; Stress Management, CAM Approach.

Further Reading

Barr-Taylor, C. B., Houston-Miller, N., Smith, P. M., et al. (1997). The effects of a home-based, case-managed, multifactorial risk-reduction program on reducing psychological distress in patients with cardiovascular disease. *Journal of Cardiopulmonary Rehabilitation* **17**, 157–162.

Bennett, P. and Carroll, D. (1994). Cognitive-behavioural interventions in cardiac rehabilitation. *Journal of Psychosomatic Research* **38**, 169–182.

Berkman, L. F., Blumenthal, J., Burg, M., et al. (2003). Effects of treating depression and low perceived social support on clinical events after myocardial infarction: the Enhancing Recovery in Coronary Heart Disease Patients (ENRICHD) Randomized Trial. *Journal of the American Medical Association* **289**, 3106–3116.

Bundy, C., Carroll, D., Wallace, L., et al. (1998). Stress management and exercise training in chronic stable angina pectoris. *Psychology and Health* **13**, 147–155.

Eisenberg, D. M., Delbanco, T. L., Berkey, C. S., et al. (1993). Cognitive behavioural techniques for hypertension: are they effective? *Annals of Internal Medicine* **118**, 964–972.

Frasure-Smith, N., Lesperance, F., Prince, R. H., et al. (1997). Randomised trial of home-based psychosocial nursing intervention for patients recovering from myocardial infarction. *Lancet* **350**, 473–479.

Johnston, D. W., Gold, A., Kentish, J., et al. (1993). Effect of stress management on blood pressure in mild primary hypertension. *British Medical Journal* **306**, 963–966.

Jones, D. and West, R. (1995). *Cardiac rehabilitation.* London: BMJ Publishing Group.

Jones, D. A. and West, R. R. (1996). Psychological rehabilitation after myocardial infarction: Multicentre randomised controlled trial. *British Medical Journal* **313**, 1517–1521.

Lewin, R. J. P., Furze, G., Robinson, J., Griffith, K., Wiseman, S., Pye, M. and Boyle, R. (2002). A randomised controlled trial of a self-management plan for patients with newly diagnosed angina. *British Journal of General Practice* **52**, 194–201.

Linden, W., Stossel, C. and Maurice, J. (1996). Psychosocial interventions for patients with coronary artery disease. *Archives of Internal Medicine* **156**, 745–752.

Ornish, D., Brown, S. E., Scherwitz, L. W., et al. (1990). Can lifestyle changes reverse coronary heart disease? The Lifestyle Trial. *Lancet* **336**, 129–133.

Rees, K., Bennett, P., West, R., et al. (2004). Psychological interventions for coronary heart disease. *The Cochrane Database of Systematic Reviews* **3**, CD002902.pub2.

Stress Management, CAM Approach

K Krebs
Allegheny General Hospital, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 532–537, © 2000, Elsevier Inc.

Mind/Body Medicine Strategies
Bioenergy Medicine Strategies
Manual Healing Strategies
Nutrition, Botanical Medicine and Nutritional Supplements

Glossary

Complementary and alternative medicine (CAM)	The use and practice of therapies and diagnostic techniques that may not be part of any current Western health-care system, culture, or society. The emphasis is on prevention, treatment, or both, using the self-healing capacity within a mind/body approach (in certain but not all cases).

Mind/Body Medicine Strategies

Mind/body medicine is an umbrella term that covers activities and therapies that focus on the self-healing potential within us through the interrelationship of mind and body. The goal in mind/body medicine is to optimize health by combining mental, physical, emotional, and spiritual approaches, which in turn influence the hormonal, nervous, and immune systems. Yoga, for example, involves both physical movement and a meditative state of mind and may serve to enhance a person's physical stamina and endurance while diminishing the effects of stress and perhaps strengthening one's personal relationship with God or a higher power.

Meditation and Yoga

Meditation can have significant effects on cardiovascular health. In a study conducted at an inner-city health care center, older people with mild hypertension were divided into three groups: those who learned transcendental meditation (TM), a concentration-meditation practice that involves focusing on a single object such as a sound, a word, or short phrase; those who participated in a progressive muscle-relaxation program; and those who were educated in lifestyle changes. At 3 months follow-up, the TM and muscle relaxation groups fared the best in terms of lowered blood pressure, with TM proving to be twice as effective as muscle relaxation.

Using the relaxation response, a simple meditative technique in which the participant sits quietly, focusing on breathing or a word, Harvard's Mind/Body Medical Institute researchers found that 80% of hypertensive patients had lowered blood pressure and decreased medication needs and that 16% were able to discontinue all their antihypertensive medications. These results persisted throughout the 3-year follow-up period. French researchers tested 50 hypertensive people as they sat silently, read, or talked about stress in their lives. Those who talked about stress experienced a sharp increase in systolic blood pressure, an average of 17 mm, an indication that substantiates

further how the mind (thoughts, feelings, and perceptions) can influence our physiology.

Numerous studies have shown that stress is a major trigger for flare-ups of psoriasis, and a study by British researchers suggested that the onset of psoriasis may be influenced by stress as well. Of 64 people with chronic skin problems, including psoriasis, 44 of the participants had experienced a major life event, such as a serious illness or a death in the family, soon before or at about the time their skin eruptions began. A small Johns Hopkins study found that some psoriasis patients who underwent hypnosis and visualized their condition clearing up on its own enjoyed a complete remission of symptoms. In a study of 51 chronic pain sufferers, ranging from back pain to chest pain (angina), a form of mindfulness meditation was used in a 10-week program. Sixty-five percent of patients reported a significant reduction in their perceived pain and 50% said their pain was reduced by more than half. Significant reductions in mood disturbances and psychiatric symptoms accompanied the reduction in pain as well.

Harvard's Mind/Body Medical Institute found that chronic pain patients reduced their physician visits by 36% following implementation of the regular practice of the relaxation response. Equally impressive is that this same technique has been credited for 75% of insomnia patients becoming normal sleepers and 90% having reduced or entirely eliminated sleeping medications.

Guided Imagery

Guided imagery teaches people to imagine scenarios that may help influence certain physiological conditions. Headache sufferers have successfully utilized their imaginations to release the vise grip constraining their face and scalp. Cancer patients have imagined their immune system partnering with their chemotherapy to escort out malignant cells through the lymphatic system. Several pilot studies have demonstrated elevated natural killer cells and elevated levels of serum and salivary immunoglobulin levels following guided imagery interventions. Sufferers of irritable bowel syndrome are taught to envision their digestive tract as healthy and rhythmically producing normal peristaltic contractions to create a harmony between mind and gut and normal evacuation. The University of Arizona's Integrative Medicine/Children's Research Center studies complementary and alternative medicine (CAM) methods of pediatric disorders for which there seems to be no conventional medical therapies available. One study tests the use of relaxation, guided imagery, and chamomile tea in children with recurrent abdominal pain.

Prayer and Spirituality

Prayer and spirituality are recognized CAM therapies by the NIH and funded studies are underway. The University of Michigan is studying the effects of spirituality and religiosity following cardiac surgery. A review of the psychiatric literature revealed 87% of the research that has examined religious involvement as a variable reports a positive effect, with the remaining 13% reporting a negative impact on mental health.

More than 130 controlled laboratory studies show that prayer, or a prayer-like state of compassion, empathy, and love, can bring about healthful changes in many types of living things, from humans to bacteria. This does not mean that prayer always works, any more than drugs and surgery always work, but that, statistically speaking, prayer is effective. Faith helps mobilize a person's defenses and assists in getting well, and optimism generally leads to better outcomes. Many case histories and scientific studies affirm this observation. Oxman and colleagues investigated the role religious feeling and activity played in 232 patients over the age of 55 undergoing cardiac surgery. Those who derived at least some strength, comfort, and hope from religion were more likely to survive longer after cardiac surgery than those who did not. Numerous studies in humans show that we can die as a result of dire beliefs and a sense of overwhelming futility.

Aromatherapy

Aromatherapy is the use of pure essential oils to promote health. Smell is one of the most basic human senses that can evoke a variety of emotions. Aromatherapy can be used in baths, as a diffusion or inhalation, or in massage therapy. Relieving anxiety, reducing stress, and promoting relaxation are perhaps the most well known of the uses for aromatic essential oils. Chamomile is best known for its calming properties. Lemon balm and orange blossom can have profound effects on the reduction of depression and anxiety. Lavender is said to promote a peaceful mood and create a sense of calm. Research around the globe is validating aromatherapy as a medicinal modality. A double-blind, placebo-controlled study in Germany found that oil of peppermint applied to the temples reduced tension and headache pain, perhaps by increasing blood flow to the skin. Aromatherapists believe that the fragrance of oils has a soothing effect on the brain's limbic system, which is involved in memory, emotion, and hormone regulation. Substances that affect mood also affect brain chemistry and the immune system. For example, the essential oil of ylang-ylang, a tropical flower, has been shown to cause the pituitary to

secrete more euphoric endorphins. The scent of oil of marjoram boosts the production of calming serotonin, and oil of grapefruit stimulates the brain to produce natural painkillers called enkephalins. However, critics of aromatherapy suggest that the relaxation attributed to the oils may actually be due to their application by massage, hot baths, the human touch, and other pleasant methods.

Bioenergy Medicine Strategies

This category of complementary modalities is based on the concept that the human energy field is subject to imbalances and that restoration of the energy fields improves health and sleep and promotes relaxation.

Acupuncture

In November 1997, a National Institute of Health panel concluded that acupuncture is an acceptable alternative or adjunct to conventional therapy for headaches, low back pain, fibromyalgia, asthma, and carpal tunnel syndrome. The panel stressed that acupuncture has minimal side effects compared to conventional treatments and noted considerable evidence that acupuncture causes the release of natural pain-relieving substances (endorphins) in the brain. In 1993, an NIH-funded study evaluated the use of acupuncture in the treatment of unipolar depression in 33 women. Based on the results of the small sample, acupuncture appeared to offer significant symptom relief comparable to conventional interventions such as psychotherapy or pharmacotherapy. The use of acupuncture as an effective treatment for depression was substantial enough to warrant a larger-scale clinical trial. Beginning in 1999, the Minneapolis Center for Addiction and Alternative Medicine began exploring the use of electroacupuncture as it pertains to substance abuse, beginning with nicotine and progressing to such substances as heroin and cocaine. Similarly, acupuncture research is being conducted at the University of Arizona. The study will evaluate the use of self-hypnosis, acupuncture, and osteopathic manipulations for children with spastic cerebral palsy.

Therapeutic Touch

Therapeutic touch (TT) is a body energy-field technique in which hands are passed over the body without actually touching to recreate and change energy imbalances for restoring innate healing forces. One of the first studies to be funded by the Office of Alternative Medicine at the NIH monitored the effects of TT on 20 very stressed nursing and medical students facing professional board exams. Students were randomized into two groups: one receiving TT and one receiving no TT. Differences between groups were found to be significant for IgA, IgM, and apotosis. Theory building is still needed to explain results that show changes in the immune response to stress-reducing therapies. This study was determined useful in documenting that measurable changes do occur following TT. In a 1986 study on 60 patients with tension headaches, patients who received TT reported an average pain reduction of 70%. The control group receiving mock TT reported an average pain reduction of 37%. Open-heart surgery patients who received TT had lower blood pressure than those who did not. While TT has gained its share of skeptics, the technique is now practiced by an estimated 30 000 health professionals around the world, is offered in at least 200 U.S. hospitals, and is taught in more than 80 universities.

Reiki

Reiki comes from the Japanese word meaning universal life force energy. The practitioner serves as a conduit for healing energy directed into the body or energy field of the recipient with or without physical contact with the body. The vast majority of reports on the effectiveness of reiki appear in popular rather than scientific literature; however, one study showed that reiki can increase hemoglobin and hematocrit levels. The study compared 47 people participating in reiki training and a small control group of 9 healthy medical professionals. The study found a significant increase in the hemoglobin and hematocrit levels among the reiki group and no significant change in the control group.

Manual Healing Strategies

Manual healing strategies is the term used for many body work techniques from both ancient and modern traditions that promote relaxation and treat ailments through reeducation in posture and movement and various forms of bodily manipulations. Acupressure, osteopathy, reflexology, and rolfing, a deep tissue massage used to release fascial adhesions, are just a few of the manual healing strategies that can aid in managing stress-related symptoms.

Massage

Massage therapy includes an array of manual therapies that manipulate the soft tissues of the body in order to reduce anxiety, tension, stress, and depression. In addition, massage therapy increases circulation, controls pain, and promotes a sense of overall well-being. Touch is the main ingredient of massage therapy, which conveys a sense of caring – an important component in the healing partnership. A study of the use of massage in patients with pain from cancer found that those who received massage had a 60% lower level of pain perception and a 24% decrease in

anxiety. More than half reported feeling more relaxed as well. In a study of children and adolescents hospitalized with depression and adjustment disorder, those who were treated with massage were found to be less anxious and depressed than those who viewed relaxation videotapes.

Nutrition, Botanical Medicine and Nutritional Supplements

A healthy diet can help the body's healing system operate more efficiently, as has been demonstrated in a number of clinical research studies. In the 1970s, Dr. Dean Ornish, a medical internist, took the medical community by storm when he suggested that a vegetarian diet (10% fat), combined with yoga, meditation, moderate exercise, and group support, could reverse heart disease. Now clinical studies have demonstrated that once-occluded coronary vessels are functioning optimally, and people subjectively report a return to a healthier, happier lifestyle free from anginal pain. Many third-party payers now cover participation in this lifestyle program to attempt to avoid the more costly coronary bypass surgery and coronary angioplasty. Comparisons of a Western and Eastern lifestyle program in healthy adult subjects showed that those adhering to an Ayurvedic (Eastern) lifestyle similar to the Dean Ornish program visited their physicians less frequently, used fewer prescription medications, and had less episodes of depression than the Western lifestyle group.

Sales of soy-based products such as soy milk, tofu, and tempeh have skyrocketed as a result of reports of their protective effects against breast and prostate cancer and their ability to minimize perimenopausal symptoms. Organic produce rich in natural antioxidants are said to have a positive effect on the nervous system, thus aiding in stress reduction. Coffee substitutes and herbal teas are replacing caffeinated beverages, which contribute to nervousness, anxiety, palpitations, and insomnia. Botanical medicine sales aimed at stress reduction have soared amid claims of safety and less side effects than their pharmaceutical counterparts. The German Commission E, equivalent in many ways to the Food and Drug Administration in the United States, has conducted clinical trials on several botanical medicines for anxiety and mild depression. Kava is derived from the knotty root of a large tropical shrub in the black pepper family. This herb has been widely used in Europe for nervous anxiety, stress, and sleep aid.

Research has demonstrated that kavalactones are responsible for the induced relaxation that acts on the central nervous system. In 1996 in Germany, a double-blind randomized trial investigated the effects of kava on 58 adults with anxiety. Half the subjects took 100 mg of kava extract (standardized to 70 mg kavalactones) three times per day for a month; the other group took a placebo. After 1 week the kava group showed significant improvement in symptoms compared with the placebo group and continued to improve over the course of the research study. Kava is recommended for short-term use and should not be taken with alcohol, prescription sedatives, or valerian root, a herbal sedative. Hypericum perforatum, better known as St. John's wort, is a shrubby perennial plant that, when used in a standardized extract containing the main active component hypericin, improves mood and sleep in those suffering from mild to moderate depression. German researchers are currently studying hypericin, another component of St. John's wort thought to be superior in alleviating mild to moderate depression. The recommended dose is 300 mg three times a day and usually requires at least a month of regular use to achieve maximum symptom management. Following an animal study suggesting that St. John's wort may inhibit alcohol cravings by influencing levels of serotonin and other neurotransmitters, researchers at the University of North Carolina at Chapel Hill are planning a study using the herb on human alcoholics.

Insomnia is one of the most frustrating symptoms associated with stress. Valerian (*Valeriana officinalis*) is a European plant that eases stress and mildly sedates the central nervous system. Valerian binds to the same receptors in the brain as prescription sleeping aids. Even though it is mild compared to pharmaceutical sedatives, valerian is a depressant and should be reserved for occasional use rather than taken nightly. Alcohol and other depressants should be avoided when using valerian root. Chamomile tea is one of the most effective bedtime beverages. This herb has a sedative effect that can be used by all ages. In addition to easing away tensions to help with sleep, chamomile soothes indigestion or heartburn and, for infants, is a helpful remedy for colic.

In addition, there are supplements that aid naturally in the brightening one's mood. Dr. Andrew Weil, from the University of Arizona's Integrative Medicine Program, recommends a simple formula intended to change the balance of neurochemicals in the brain to support energy and mood elevation. In the morning, upon awakening, the amino acid DL-phenylalanine (DLPA) is taken. If sufficient mood elevation has not been achieved after trying DLPA, then L-tyrosine, another amino acid, can be substituted. One may switch back and forth between these amino acids. In addition, vitamins C and B_6, a glass of fruit juice, and a small piece of fruit should be consumed and breakfast eaten 45 min later. Vitamins C and B_6 are

repeated in the evening. These natural supplements appear to be safe to take, even if one is currently on selective serotonin reuptake inhibitor antidepressants. Dr. Weil asserts that aerobic activity is actually the best antidepressant because it increases the production of endorphins, the brain's own opiate-like molecules that are associated with a natural high.

See Also the Following Articles

Integrative Medicine (Complementary and Alternative Medicine); Hypnosis; Relaxation Techniques; Religion and Stress.

Further Reading

Benson, H. (1975). *The relaxation response*. New York: Avon Publisher.

Benson, H. (1996). *Timeless healing: the power and biology of belief*. New York: Scribner.

Chevallier, A. (1996). *The encyclopedia of medicinal plants*. London: Dorling Kindersley.

Dossey, L. (1993). *Healing words and prayer is good medicine*. New York: Harper.

Duke, J. (1997). *The herbal pharmacy*. Rodale Press.

Ferrell-Torry, A. T. and Glick, O. J. (1993). The use of therapeutic massage as a nursing intervention to modify anxiety and the perception of cancer pain. *Cancer Nursing* **16**, 93–101.

Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation. *General Hospital Psychiatry* **4**, 33–47.

Keller, E. and Bzdek, V. M. (1986). Effects of therapeutic touch on tension headache pain. *Nursing Research* **35**, 101–105.

Larson, D. (1996). *The role of prayer in mental health*. Alternative Therapies.

Leake, R. and Broderick, J. (1998). *Treatment efficacy of acupuncture: a review of the research literature. Integrative Medicine* **I** (Issue 3).

National Institutes of Health (NIH) (1994). *Alternative medicine: expanding medical horizons, a Report to the National Institute of Health on Alternative Medical Systems and Practices in the United States*. p. 143. NIH Publication No. 94–066.

Olson, M. (1993). *Immune response to therapeutic touch during stress*. NIH: Alternative Medicine Grant.

Ornish, D. (1998). *Love and survival: the scientific basis for the healing power of intimacy*. New York: Harper Collins.

Oxman, T. E., Freeman, D. H., Jr. and Manheimer, E. D. (1995). Lack of social participation or religious strength and comfort as risk factors for death after cardiac surgery in the elderly. *Psychosomatic Medicine* **57**, 5–15.

Schneider, R. H., et al. (1995). A randomized controlled trial of stress reduction for hypertension in older African Americans. *Hypertension* **17**, 192–200.

Simon, D. (1993). *Ayurvedic Medicine Demonstration Project*. NIH: Alternative Medicine Grant.

Spencer, J., et al. (1999). *Complementary/Alternative medicine: an evidence-based approach*. New York: Mosby.

Weil, A. (1995). *Health and healing*. New York: Houghton Mifflin.

Weil, A. (1995). *Natural health, natural medicine*. New York: Houghton Mifflin.

Weil, A. (1995). *Spontaneous healing: how to discover and enhance your body's natural ability to maintain and heal itself*. New York: Knopf.

Weil, A. (1996). *Self healing: creating natural health for your body and mind*. New York: Houghton Mifflin.

Wetzel, W. (1989). Reiki healing: a physiologic perspective. *Journal of Holistic Nursing* **7**(1).

Stress of Self Esteem

J C Pruessner, S Wuethrich and M W Baldwin
McGill University, Montreal, Quebec, Canada

© 2007 Elsevier Inc. All rights reserved.

The Role of Personality in the Perception of Stress

The Importance of Self-esteem and Locus of Control in the Perception of Stress

Endocrinological Evidence for the Role of Self-esteem and Locus of Control in the Perception of Stress

The Hippocampus as a Possible Mediator of the Relationship between Self-esteem, Locus of Control, and Stress

Conclusion: Personality Variables, Brain Structures, and Stress Responses in a Developmental Context

Glossary

Cortisol	A steroid hormone secreted by the human adrenal glands. It is often released in high amounts during periods of stress. High doses lead to interference with the proper functioning of the immune system.
Hippocampus	Seahorse-shaped structure that extends along the inferior horn of each lateral ventricle of the brain and consists of gray matter covered on the ventricular

	surface with white matter. A part of the limbic system, it is involved in the storage of memory for intermediate periods and the consolidation of those memories into permanent form.
Hypothalamic-pituitary-adrenal (HPA) axis	The major part of the neuroendocrine system that controls reactions to stress and has important functions in regulating various body processes such as digestion, the immune system, and energy usage. Most species (from humans to simple organisms) share components of the HPA axis. It regulates a set of interactions among glands, hormones, and parts of the midbrain that mediate the general adaptation syndrome proposed by Selye.
Locus of control	The perception of factors responsible for the outcome of an event. An individual with high internal locus of control believes that his or her own actions caused the outcome. Conversely, an individual with high external locus of control believes the outcome was determined by outside forces.
Personality	The complex of all attributes – behavioral, temperamental, emotional, and mental – that characterize an individual.
Self-esteem	The global value we place on ourselves.
Stressor	Any emotional or physical demand (positive or negative) that gives or results in stress.

The Role of Personality in the Perception of Stress

Lazarus, who has devoted much of his work to the study of stress, noted early on that stressful conditions per se fail to produce reliable effects on task performance. Keeping all situational variables constant, the same stressor can have minimal effect, lead to performance improvement, or result in significant performance impairment across subjects. This led him and others to suggest that individual differences in motivational and cognitive variables are likely to interact with situational components, and what is considered threatening for some, and thus stressful and performance impairing, might be considered as stimulating by others, and thus produce beneficial effects on performance.

Subsequent research identified the process of evaluation as crucial in explaining the impact of psychological variables on the stress response. Any internal or external stimulus is perceived as stressful only if it is evaluated as harmful or threatening. The early notion of the importance of the evaluation has received recent validation with the identification of social evaluative threat as the single most important factor in determining the stressfulness of a situation in laboratory stress studies.

The Importance of Self-esteem and Locus of Control in the Perception of Stress

A case can be made that the personality variables self-esteem and locus of control play a central role in the evaluation of many situations and thus contribute to the experience of stress. Self-esteem is broadly defined as the value we place on ourselves. Epidemiological studies have shown that low self-esteem is associated with negative life outcomes, including substance abuse, delinquency, unhappiness, depression, and worsened recovery after illnesses. On the other hand, high self-esteem has been linked to happiness and longevity. In studies of aging, a positive self-concept has been identified to play a key role in successful aging, predicting independence, cognitive stability, and general health in old age.

Internal locus of control can be defined as the perception that one's outcomes are determined by one's actions. Not surprisingly, internal locus of control and self-esteem are usually highly correlated. The key link of these variables to the experience of stress lies in their impact on the evaluation of any given situation. We postulate that in the evaluation of whether a given situation is threatening or benign, self-esteem and internal locus of control systematically interact with situational factors. If a person attributes little importance to him- or herself and thinks that he or she has little impact on the outcome of his or her own actions, this person will find more situations uncontrollable and unpredictable, and consequently will evaluate more situations as threatening and harmful.

Endocrinological Evidence for the Role of Self-esteem and Locus of Control in the Perception of Stress

Evidence for the impact of self-esteem and locus of control on stress perception emerged when subjects were exposed to repeated psychological stress, using the Trier Psychosocial Stress Test (TSST). In this paradigm, subjects have to give an impromptu speech and perform serial subtraction tasks in front of an audience, usually for about 10 min. The audience consists of two to three persons who are instructed to maintain a neutral expression, being neither explicitly rejecting nor confirmative in their facial expression or gestures. During the speech, the audience interacts with the subject only to indicate the amount of time that is left to talk or to ask specific questions.

In the case that a subject stumbles, they encourage the subject to continue the speech. During the serial subtraction task, the subject is interrupted only when making a mistake. The subject is then corrected and instructed to start the task over. The task was designed to represent a significant social-evaluative threat and indeed has been shown to be a powerful stressor, stimulating the hypothalamic-pituitary-adrenal (HPA) axis and leading to significant free cortisol increases within 15 to 30 min following the onset of the task. This first study aimed to validate the long-standing hypothesis that in humans, repeated exposure to the same stressor would lead to quick habituation of the stress response. In order to test the habituation of the stress response, 20 young healthy male college students were exposed to the TSST on five subsequent days. For this purpose, the TSST was modified using different speech topics and serial subtraction tasks on each day. Interestingly, only 13 of the 20 subjects showed the typical habituation pattern, with a normal stress response on day one being significantly reduced on day two, and no longer present on the subsequent days. In the seven remaining subjects, however, the cortisol stress response continued to be present on all days and only showed a tendency to decline toward the end of the testing (**Figure 1**). When analyzing the available psychological variables, it became apparent that low internal locus of control and low self-esteem were the best predictors of failing habituation of the cortisol stress response to repeated stress exposure. This can be interpreted as a sign that these personality variables interact with the evaluation of a situation during repeated exposure. The absence of differences in the stress response between the two groups of subjects on day one was at the time attributed to the effect of novelty – the novelty of the situation might have made it unpredictable and uncontrollable for everybody on the first exposure and might thus have masked the impact of personality variables on stress perception and response. One conclusion at the time was that in order to reveal the effect of personality variables on the stress evaluation and response, one would likely need repeated exposures to the same stressor in order to reveal the influence of personality variables on stress.

However, it is known that personality variables tend to have relatively weaker effects when situational factors are very strong. In a second study, the threatening aspects of the situation were reduced, and self-esteem and locus of control had an impact on the perception of stress on the first exposure to a stimulus. Here, computerized mental arithmetic was combined with an induced failure design to invoke stress. In the setup used in this task, 52 students performed the task on computer terminals in front of them. Half of the students were exposed to a difficult version leading to low performance, compared to an easy version of the task with high performance for the other half. The students played the task in three 3-min segments and had to announce their performance score after each segment to the investigator, who wrote the scores down on a board for everybody to see. Saliva sampling before, throughout, and after the task allowed the assessment of the cortisol dynamics in relation to this paradigm. Interestingly, this task triggered a significant cortisol release only in the subjects who were in the low-performance group and had low self-esteem and low internal locus of control. Neither low performance alone nor low self-esteem and internal locus of control alone were significant predictors of cortisol

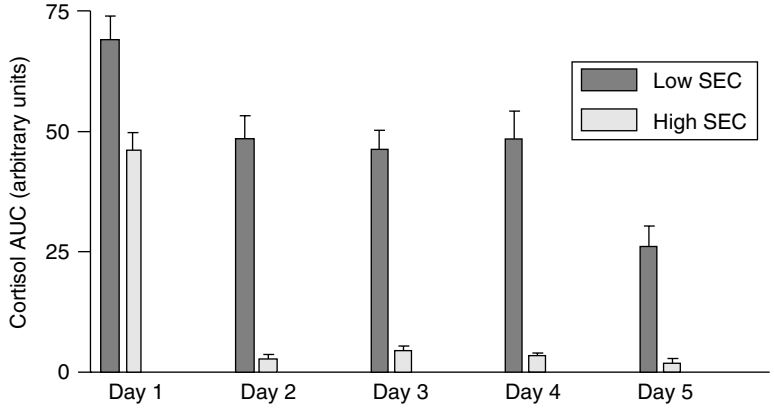


Figure 1 Cortisol responses (AUC, area under the curve) on repeated exposure to the Trier Social Stress Test (TSST) on 5 subsequent days in subjects with high self-esteem and high locus of control (high SEC; $n = 13$) and low self-esteem and low locus of control (low SEC; $n = 7$).

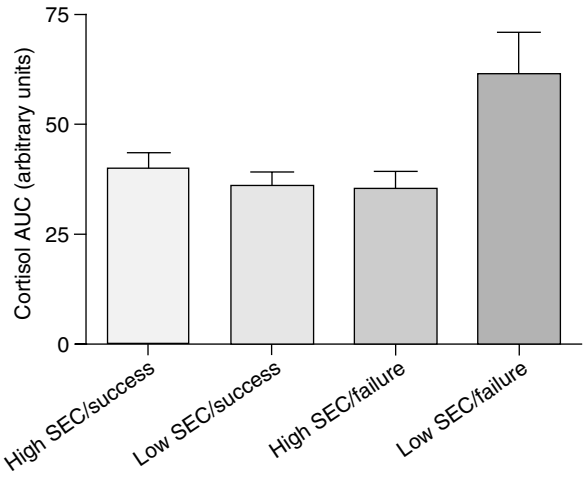


Figure 2 Cortisol stress responses to the Trier Mental Challenge Task (TMCT) in four groups of subjects, separated for high and low self-esteem and locus of control and high and low performance in the mental arithmetic. The performance was manipulated by the investigator.

release, supporting the notion that these personality variables produce effects only in interaction with a potentially stressful situation (Figure 2). The evaluation of the situation is suggested to be at the core of this interaction.

The Hippocampus as a Possible Mediator of the Relationship between Self-esteem, Locus of Control, and Stress

Although studies on brain correlates of personality variables and endocrine function in humans are only starting to appear, a picture is emerging that puts the spotlight on the hippocampus as a likely mediator between the previously reported personality variations and stress responses. The hippocampus is one of the major limbic system structures involved in the regulation of the stress response, and variations in hippocampal volume (HCV) have been postulated and shown to be systematically linked to excessive HPA axis activity. Early models postulated that associations between HCV and HPA axis activity might represent the consequence of excessive exposure of the hippocampus to glucocorticoids, due to their powerful neurotoxic properties.

More recently, however, evidence for a link between naturally occurring variations in HCV and HPA function has surfaced. Functionally, the hippocampus is the primary structure for memory contextualization, and it is here that a link to self-esteem and locus of control could occur. When an individual is faced with a potentially threatening and harmful situation, he or she employs memories of past events to contribute to the evaluation of the current event. However, if specific situational and environmental characteristics associated with negative past events cannot be recalled, this lack of awareness of situational circumstance can lead to an overgeneralization of negative past events and thus an increased likelihood to consider the current situations as stressful as well. Thus, poor contextualization mediated by the hippocampus could be linked to higher stress responses, on the one hand, and lower self-esteem on the other.

This hypothesis of a link between HCV, personality variables, and stress responses was tested in 17 healthy young male college students. All subjects underwent extensive psychological assessment for assessment of personality variables, including self-esteem and locus of control. In addition, they underwent magnetic resonance imaging employing a structural T1-weighted acquisition protocol, which produces high-resolution images of the cerebrum. A recently developed manual segmentation protocol was then applied to all images that included nonuniformity correction, signal intensity normalization, and Tailarach-like transformation for standardization of brain size. This protocol allows for manual assessment of hippocampal volumes. Finally, all subjects were exposed to a computerized stress task, similar to the mental arithmetic task described earlier. Saliva samples before, during, and after the task accompanied the testing to capture the cortisol stress response. Results first replicated earlier findings of mental arithmetic stress: only the subjects with low self-esteem and low internal locus of control showed a significant release of cortisol. Extending earlier findings, however, a correlation emerged between the cortisol stress response and self-esteem ($r = -0.45$, $p < 0.05$). Furthermore, a significant correlation emerged between the cortisol stress response and the total hippocampal volume ($r = -0.53$, $p = 0.03$), suggesting that the size of the hippocampus is related to the regulation of the HPA (Figure 3a). Finally, supporting the initial hypothesis, a link was observed between internal locus of control and HCV ($r = 0.66$, $p = 0.006$; Figure 3b) and self-esteem and HCV ($r = 0.58$, $p = 0.02$; Figure 3c).

Testing for the specificity of the effect with regard to the hippocampus, total brain gray matter was employed as control structure in correlations with both personality variables and cortisol response. None of the tested correlations were significant, suggesting that the observed relationships were indeed specific to this medial temporal lobe area.

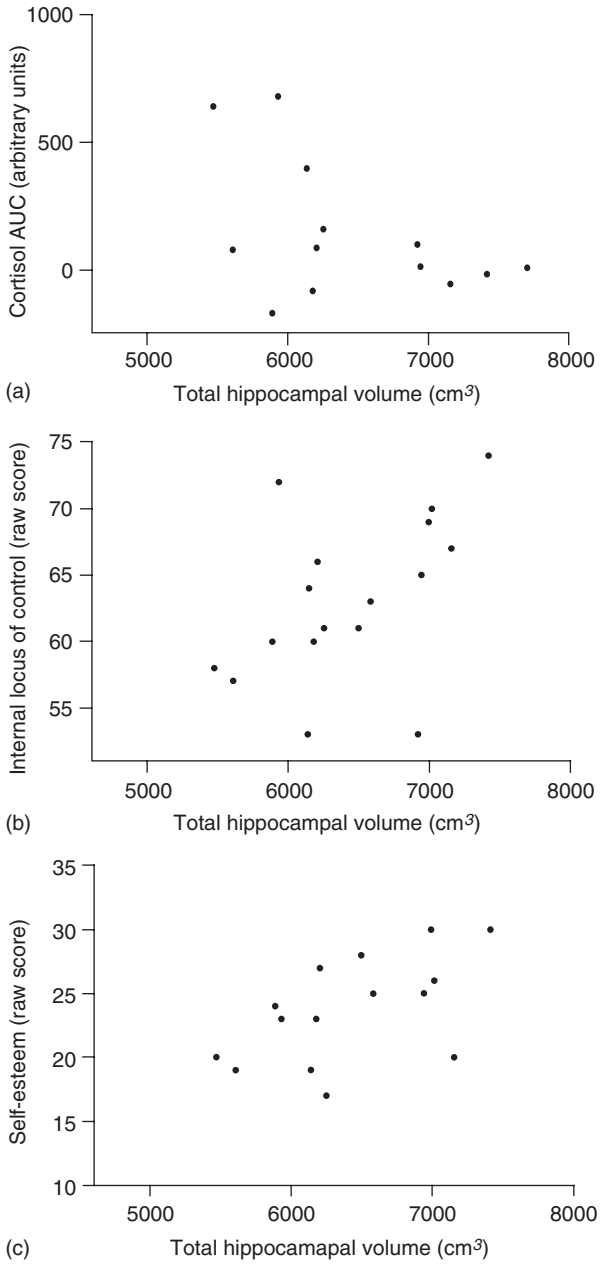


Figure 3 Correlation between hippocampal volume and cortisol stress responses to a mental arithmetic task (a), hippocampal volume and locus of control (b), and hippocampal volume and self-esteem (c) in a group of 17 healthy young male college students.

Conclusion: Personality Variables, Brain Structures, and Stress Responses in a Developmental Context

It will be a challenging task for future studies to ascertain the origin of the relationship observed between personality variables, hippocampal volume, and stress response. There are a number of possibilities as to how these associations could develop. One model puts naturally occurring variations in hippocampal volume at the origin of these functional relationships. It can be speculated that HCV is a consequence of multiple interacting factors during critical developmental stages, leading to the variation in HCV that follows a normal distribution pattern. If, however, the size of this structure determines memory contextualization capabilities, this variation would then have personality and stress-related consequences. The impaired source monitoring would lead to an overgeneralization of the experience of failure and rejection, and consequently to a self-perception of being a failure or being socially rejected in general, which might produce low self-esteem and low locus of control during childhood and adolescence. The observed higher cortisol responses would then not be a direct consequence of small HCV, but of the increased stress perception due to consistent unfavorable evaluations of ambiguous situations. Of course, the question arises why overgeneralization should then not also be an issue for positive life events, in other words, a general lack of source monitoring for both positive and negative life events with no consequences on self-perception. One possible explanation could be seen in the stress-specific function of the hippocampus within the limbic system. There is evidence that it is specifically the hippocampus together with the anterior cingulate that is involved in the stress response. It is also known that the anterior cingulate is particularly involved in error monitoring, which means that its involvement is restricted to situations in which a mismatch between expectations and reality occurs. These circumstances would make a specific role of the hippocampus in reaction to negative life events more probable.

Based on the findings to date, we suggest that reduced HCV might play a role in the development of low self-esteem. This may in turn produce a higher susceptibility of perceiving ambiguous situations as threatening, and thus stressful. The idea that HCV variation might be the cause for adverse functional and behavioral consequences, rather than their consequence, has recently been proposed in conjunction with risk factors for developing posttraumatic stress disorder (PTSD) in a study investigating hippocampal volume in twin brothers. Here, in subjects who developed PTSD as a consequence of participating in the Vietnam War, the researchers observed lower hippocampal volume compared to subjects who went to war but did not develop PTSD. Intriguingly, however, lower hippocampal volume could also be observed in the PTSD subjects' twin brothers – who never went to war – leading to the speculation that HCV might be a risk factor for, rather than a consequence of, developing PTSD.

As mentioned earlier, the possibility that HCV plays a causal role is supported by the fact that consistent results were found even among relatively young adults. Recently, variations in HCV in young populations have been reported, suggesting that there is a considerable range of hippocampal volumes already in young subjects. At the same time, self-esteem is known to be a stable trait, with considerable intra-individual stability throughout life. This supports a model in which variations in brain morphology could become a pathway to certain personality characteristics and behavioral phenotypes early in life, but longitudinal studies will be needed to examine these changes over time and provide validation for the assumed associations.

See Also the Following Articles

Corticosteroid Receptors; Hippocampus, Corticosteroid Effects on; Hippocampus, Overview; Hypothalamic-Pituitary-Adrenal; Optimism, Pessimism, and Stress; Psychological Stressors, Overview; Social Status and Stress.

Further Reading

Baltes, M. M. and Baltes, P. B. (1990). *Successful aging*. Cambridge, UK: Cambridge University Press.

Baumeister, R. F., Campbell, J. K., Krueger, J. I. and Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest* **4**, 1–44.

Dickerson, S. S. and Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin* **130**(3), 355–391.

Fuchs, E. and Flugge, G. (2003). Chronic social stress: effects on limbic brain structures. *Physiology and Behavior* **79**(3), 417–427.

Gilbertson, M. W., Shenton, M. E., Ciszewski, A., Kasai, K., Lasko, N. B., Orr, S. P., et al. (2002). Smaller hippocampal volume predicts pathologic vulnerability to psychological trauma. *Nature Neuroscience* **5**(11), 1242–1247.

Hoyle, R., Kernis, M. H., Leary, M. R. and Baldwin, M. W. (1999). *Selfhood: identity, esteem, control*. Boulder, CO: Westwood.

Kirschbaum, C., Bartussek, D. and Strasburger, C. J. (1992). Cortisol responses to psychological stress and correlations with personality traits. *Personality and Individual Differences* **13**(12), 1353–1357.

Lazarus, R. S. (1999). *Stress and emotion: a new synthesis*. New York: Springer.

Leary, M. R. and McDonald, G. (2003). Individual differences in self-esteem: a review and theoretical integration. In: Leary, M. R. & Tangney, J. P. (eds.) *Handbook of self and identity*, pp. 401–420. New York: Guilford Press.

Petrie, K. and Rotheram, M. J. (1982). Insulators against stress: self-esteem and assertiveness. *Psychological Report* **50**(3 Pt 1), 963–966.

Pruessner, J. C., Hellhammer, D. H. and Kirschbaum, C. (1999). Low self-esteem, induced failure and the adrenocortical stress response. *Personality and Individual Differences* **27**(3), 477–489.

Pruessner, J. C., Collins, D. L., Pruessner, M. and Evans, A. C. (2001). Age and gender predict volume decline in the anterior and posterior hippocampus in early adulthood. *Journal of Neuroscience* **21**(1), 194–200.

Pruessner, J. C., Baldwin, M. W., Dedovic, K., et al. (2005). Self-esteem, locus of control, hippocampal volume, and cortisol regulation in young and old adulthood. *Neuroimage* **28**, 815–826.

Sapolsky, R. M. (1999). Glucocorticoids, stress, and their adverse neurological effects: relevance to aging. *Experimental Gerontology* **34**(6), 721–732.

Sinha, R., Lacadie, C., Skudlarski, P. and Wexler, B. E. (2004). Neural circuits underlying emotional distress in humans. *Annals of the New York Academy of Sciences* **1032**, 254–257.

Stress System Balance Hypothesis

E R de Kloet
Leiden/Amsterdam Center for Drug Research and Leiden University Medical Center, Leiden, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

Corticotropin Releasing Hormone and Urocortins
Glucocorticoids
Two Corticosteroid Receptor Types: Low and High Levels of Stress System Operation
Corticosteroid Receptor Balance
Nuclear Receptors: Slow Genomic and Fast Membrane Actions
Perspectives

Glossary

Allostasis	The ability to reestablish homeostasis through change during adaptation to stress.
Allostatic load	The cost of allostasis; a too high cost may lead to wear and tear, and ultimately to disease.
Corticotropin releasing hormone (CRH)	A hormone produced in the hypothalamus that stimulates after secretion in the portal vessels the release of pituitary ACTH.
Homeostasis	A stable *milieu intérieur*, which refers to the defense of physiological and psychological integrity.
Homeostatic mechanism	The processes and activities coordinated by the stress system that help to maintain homeostasis.
Hypothalamic-pituitary-adrenal (HPA) axis	Refers to a coordinate pattern of stimulation and feedback interaction between hypothalamic paraventricular neurons, pituitary corticotrophs and the adrenal cortex with the aim to faciliate adaptation to stress and to synchronize daily and sleep-related events.
Nuclear receptors	Proteins that bind steroid hormones and that can act in the cell nucleus as gene transcription factors in the modulation of gene expression. The receptors appeared recently to reside also in the membrane where they mediate fast non-genomic actions. In the brain, mineralocorticoid receptors are promiscuous and bind with high affinity naturally occurring glucocorticoids (cortisol and corticosterone), mineralocorticoids, and progesterone; glucocorticoid receptors bind exclusively naturally occurring and synthetic glucocorticoids.
Resilience	The characteristic of a stress system that rapidly turns on and off.
Secretagogues	Substances that stimulate the endocrine system to secrete hormones; here used to indicate the peptides released by hypothalamic neurons in the pituitary portal vessel system.
Set point	The set level of the regulated variable (physiological function); homeostatic mechanisms serve to restore the set point perturbed by stress. Allostasis reflects an adaptive change in the set point necessary to cope with the stress.
Stressor	Any perceived threat to the individual's integrity.
Stress system balance hypothesis	The hypothesis that the defense of homeostasis and development of allostasis depend on the balanced interaction between activating and inhibiting humoral and neural stress system mediators.
Stress system mediators	The network of humoral and nervous mediators of the stress response.
Stress response	The spectrum of physiological and behavioral adaptations that defend homeostasis and/or to achieve allostasis and that are coordinated by the stress system mediators.

A stressor is any real or perceived threat to homeostasis (i.e., the stability and integrity of the organism). The defense of homeostasis to the stressor triggers a corrective physiological and behavioral response (the stress response) that is coordinated by humoral and nervous mediators, the stress system mediators. In particular, psychological components of the stress response often require adaptation to the stressor through change, a process termed allostasis. The stress-induced activation and termination of the action of the mediators in humoral and nervous signaling represent two compensatory phases: some mediators act in an opposing mode and others in a synergistic mode (Table 1). The two modes of stress system operation need to be in balance, however, to ensure a proportionate outcome in the third and preparative phase for the next stress response. This implies the storage of the experience in memory and the storage of energy resources for coping with future needs. A disproportionate outcome of stress system operation has been termed allostatic load and is considered an unfavorable prediction for health.

Table 1 Some selected mediators of the two modes of stress system operation[a]

Activating mode: fight-flight-fright	*Suppressing mode: coping and adaptation*
CRH, CRH-1 receptors	Urocortins, CRH-2 receptors
Vasopressin	Oxytocin
ACTH	Endorphins
Sympathetic response	Parasympathetic response
Mineralocorticoid receptors	Glucocorticoid receptors

[a]The two modes are in balance for homeostasis and health. ACTH, adrenocorticotropic hormone; CRH, corticotropin releasing hormone.

Such a disproportionate outcome is reflected in an increased vulnerability to stress-related diseases. In the brain, such diseases are characterized by emotional lability and cognitive disturbance. Both may have profound consequences for metabolic, immune, and cardiovascular regulations.

Corticotropin Releasing Hormone and Urocortins

Most mediators of the stress system are known and their molecular and anatomical organization is being unraveled. Hierarchically, the hypothalamic paraventricular nucleus (PVN) integrates all stressful information. The PVN is activated directly by physical stressors through aminergic ascending pathways and by metabolic disturbance through peptidergic hypothalamic pathways. Psychologically stressful information is conveyed transsynaptically to the PVN after limbic-midbrain processing. The PVN contains different cell types that express peptidergic secretagogues and that, because of their projection to different targets, can activate distinct functions in a coordinate manner.

Hence, the PVN is the organizer of the immediate behavioral, autonomic, and neuroendocrine response to a stressor. The ratio of PVN secretagogues can change as a function of the nervous input. Corticotropin releasing hormone (CRH) is released in changing ratios with vasopressin and other peptides driving the sympathetic and behavioral fight-flight-fright response. The CRH-1 receptors mediating this response can also activate the hypothalamic-pituitary-adrenal (HPA) axis, leading to the secretion of adrenal corticosteroids. To achieve this goal, CRH and vasopressin are secreted in the pituitary portal vessel system and induce in the anterior pituitary the synthesis of proopiomelanocortin (POMC), which is processed to balanced concentrations of adrenocorticotropic hormone (ACTH), opioids, and melanocortin peptides, for example. ACTH stimulates the adrenal cortex to secrete cortisol (humans) and corticosterone (humans, rats, and mice), whereas opioids attenuate stress reactions, resulting in suppression of adrenal function. Collectively, these mediators of the HPA axis act in concert to coordinate the neuroendocrine stress response. The mechanism underlying the CRH-induced sympathetic outflow and behavioral response pattern is not precisely known.

Recently, urocortin II and III (stresscopin and stresscopin-related peptide, respectively) were identified acting through the CRH-2 receptor system. The CRH-1 and CRH-2 receptor systems have a partly overlapping distribution in the brain, which matches their CRH and urocortin terminal fields. The intracerebroventricular administration of CRH mimics the initial behavioral, autonomic, and neuroendocrine stress response and is anxiogenic. In contrast, urocortin seems to have opposing functions as expressed, for example, in its anxiolytic properties. It is well established that the CRH/CRH-1 receptor system organizes the onset of the stress response.

Some investigators have proposed that the urocortin–CRH-2 receptor system represents a slower mode of operation of the stress system mediators, perhaps one that is primarily concerned with adaptation and coping with stressors. Such a bimodal CRH-1/CRH-2 receptor-driven neuroendocrine system would control over time the compensatory phases of initiation and termination in the stress response. Such a system would operate in an integrative fashion with the sympathetic and parasympathetic divisions of the autonomic nervous system, and the signals released from the dendritic cell would control the balance between immune activation and tolerance. In fact, other neuropeptides in the stress system, such as vasopressin and oxytocin, also have opposing actions in aspects of behavioral adaptation.

Glucocorticoids

The importance of a balance in the concentration and action of stress mediators is illustrated also by the glucocorticoid hormone (cortisol or corticosterone) secreted from the adrenal cortex. The hormone operates both in the fast and slow modes of the response to stress and produces lasting effects in the preparation of future events. In concert with the other stress mediators, the hormone has two modes of operation that are complementary. First, it acts permissively and exerts a tonic influence on cellular and tissue functions. In the behavioral realm, the hormone operates in a pro-active mode, so that the organism is prepared in anticipation of upcoming events. Second, once homeostasis is disturbed, the concentration of cortisol rises in response to HPA axis activation to

promote the recovery of disturbed homeostasis. High concentrations of cortisol operate in a reactive mode, through feedback foremost, by preventing the initial stress reactions from overshooting and to become damaging in themselves.

The two complementary modes of operation of cortisol depend on the concentration of the hormone, the context and the phase of the stress response. In the low basal mode pro-active hormone actions seemingly have a household function in which the hourly pulses of the hormone over the circadian cycle provide the capability to synchronize upcoming daily and sleep-related events and to coordinate the available mental and body resources to deal with upcoming challenges. These functions are stepped up when cortisol levels rise in the stress mode. At the same time, the redistribution of resources is facilitated, enabling the individual to cope and to suppress any overreaction to a stressor in the reactive mode. Cortisol action operating in the basal and the stress mode therefore displays an enormous diversity in action that is dependent on the context. The context is a reflection of the signals provided by the other local and systemic stress mediators that converge in time and phase in the cell or tissue in which homeostasis has been challenged. In other words, it is essential that the hormones act in the right time window on the cells and circuits in the brain that have initially led to their own secretion. If they act out of context, the hormones' action becomes damaging.

Two Corticosteroid Receptor Types: Low and High Levels of Stress System Operation

The molecular basis of the dual mode of operation of cortisol is formed by two types of receptors – mineralocorticoid receptors (MRs) and glucocorticoid receptors (GRs) – that bind *in vivo* cortisol with an order of magnitude difference in affinity. However, this only occurs if the cells lack 11β-hydroxysteroid dehydrogenase type 2 (11β-HSD2), which inactivates the steroid. The receptors have a different distribution. GRs occur in virtually every cell and are expressed most abundantly in cells with a key role in coping with stress (e.g., the liver, immune system, and limbic forebrain regions). The cortisol-responsive MRs are expressed abundantly in the heart, vascular tissue, and limbic-forebrain regions such as the hippocampus. Using selective antagonists and conditional knockout strategies, researchers have studied their functions in experimental animals from gene to behavior.

MRs and GRs are nuclear receptors that mediate the slow genomic actions of cortisol. The receptors recruit patterns of coregulators and transcription factors to facilitate their dual mode of operation. In cellular studies using the hippocampus, cortisol suppresses through MR inhibitory inputs and Ca influx. The consequence is that excitatory transmission is facilitated, and Ca^{2+} availability is suppressed, which is a favorable condition for a cell. MR activation indeed protects against damage caused by ischemic insults and suppresses apoptotic pathways.

GRs are occupied in the ultradian and circadian peaks of circulating cortisol. GRs are also occupied after stress-induced rises in cortisol; thus, in hippocampal neurons, the activation of GRs some time after exposure to the stressor suppresses the stress-induced disturbances in excitability transiently raised by excitatory input. The recovery process promoted by GRs enhances a hyperpolarizing input and Ca^{2+} influx; the latter process, although beneficial in the short run, may explain why chronically elevated high glucococorticoid levels and excitotoxic challenge create unfavorable conditions for the functioning of the hippocampal neurons.

Corticosteroid Receptor Balance

From these cellular studies and from endocrine and behavioral observations, a concept has evolved in which MRs and GRs mediate the dual mode of operation of cortisol in the limbic brain to regulate in a coordinate manner the onset and termination of the response to a stressor. Because limbic areas such as the hippocampus, amygdala, and frontal cortex are the target, the action of cortisol depends on stressful information determined by spatial and temporal cues. MRs are implicated in the appraisal of the situation and in processes involved in the prediction and choice of the appropriate behavioral reaction to deal with a challenge. Such predictive responses are important for the onset of the stress response by determining the threshold or sensitivity of the stress system to challenge. Along the same line of reasoning, MRs are in a proactive mode of operation implicated in predictable ultradian and circadian changes. High levels of cortisol facilitate through GRs the termination of the stress reactions in a reactive fashion, mobilize for this purpose the required energy resources, and facilitate recovery. Due to their occupation during pulsatile and circadian peaks, GRs may function as Zeitgeber. In preparation for future events, GRs promote memory processes and the storage of energy resources. So, the two receptor types belong to interacting signaling networks that underlie adaptive processes from appraisal of novel situations and prediction of upcoming events to recovery from the challenge and storage of the experience

in memory. They participate in the control of energy metabolism, from appetite and feeding to energy disposition and storage.

Nuclear Receptors: Slow Genomic and Fast Membrane Actions

To this analysis can be added the mechanism of action of steroids in the brain via MRs and GRs. The nuclear receptors are well known as mediating the slow genomic actions of glucocorticoids. Recently it has become evident that these receptors also serve at the membrane level to mediate rapid glucocorticoid actions. A clear example is the fast MR-mediated action of corticosterone and cortisol on the release of the excitatory transmitter glutamate in the hippocampus. These fast steroid actions probably have a function in the pro-active modes of action of the steroid in anticipated perturbances of homeostasis.

Perspectives

Three questions are at the core of stress system operation and health: how is the stress system balance regulated, what are the implications of dysregulation for health and disease, and which biomarkers can reveal a balanced stress system operation?

Regulation of Stress System Balance

Regulation of the stress system balance is by definition based on the activity of the humoral and neural mediators that are activated by stress. In the case of the corticosteroids it depends on their secretion rate; bioavailability; access to the brain through multidrug-resistant P glycoprotein; the stochiometry of receptors, coregulators and other proteins and genetic factors. Genetic variants and polymorphisms can bias the response patterns. Recently, knowledge of epigenetic control of gene expression and function has grown; it has been demonstrated that stable alterations in stress system mediators may occur as a result of environmental input and early life experience. In later life, psychosocial contexts that challenge the individual's coping ability, particularly if experienced chronically, also cause enduring changes in stress system operation.

Implications for Health and Disease

Stress mediators acting either excessively or inadequately produce a phenotype vulnerable to stress-related disease in predisposed individuals. The causes of aberrant stress system operation can be environmental, experience-related, or genetic; it is reflected in the ability to cope with a stressor, and it is also affected by additional factors such as life style and diet. It is believed that an imbalance, rather than the absolute concentration of stress system mediators, compromises resilience and enhances vulnerability to disease. This condition has been termed allostatic load.

Selye championed the balance of mineralocorticoid versus glucocorticoid actions as determining pro- versus anti-inflammatory actions. Selye proposed that imbalance therefore enhanced the risk for inflammation or infection. In modern terms, in the brain the action of one single hormone – cortisol or corticosterone – depends on the balance of MR- and GR-mediated actions. Once the balance is disturbed, the individual loses the ability to maintain homeostasis when challenged by an adverse life event. This may lead to a condition of neuroendocrine dysregulation and impaired behavioral adaptation, enhancing the precipitation of stress-related disease for which the individual is genetically predisposed.

Biomarkers

Stress-related diseases, such as depression, are complex and multifactorial diseases. They likely cannot be attributed to mutations in a single gene or to a single external event; instead, they result from the concerted actions of many subtle genetic polymorphisms and from external events, the effects of which might accumulate over time. Once traumatic life events in combination with genetic disposition have engrained long-lasting changes in MR and GR signaling, a vulnerable phenotype emerges. MRs and GRs may therefore be considered as allostatic markers. The individual's vulnerability is defined by three different levels of description: the traditional clinical phenotype (emotional reactivity and personality), the functional phenotype (neuroendocrine reactivity to stimuli, MRs, GRs, and beyond), and the genotype (e.g., polymorphisms in genes that are involved in HPA signaling or monoaminergic neurotransmission). An integrated approach that analyzes patients at all three levels may provide insight into the causes of stress-related diseases and perhaps also into new targets for treatment.

Acknowledgments

Support from the Royal Netherlands Academy of Arts and Sciences and the Netherlands Organization for Scientific Research (NWO) is gratefully acknowledged.

See Also the Following Articles

Allostasis and Allostatic Load; Corticosteroid Receptors; Corticosteroids and Stress; Feedback Systems; Glucocorticoid Negative Feedback; Glucocorticoids, Effects of

Stress on; Glucocorticoids, Role in Stress; Hippocampus, Corticosteroid Effects on; Homeostasis; Hypothalamic-Pituitary-Adrenal; Stress, Definitions and Concepts of; Survivor Guilt; CRH Anxiety Circuitary in Brain; Neuropeptides, Stress-Related.

Further Reading

Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology & Behavior* **81**, 179–209.

Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. *Journal of the American Medical Association* **267**, 1244–1252.

Day, T. A. (2005). Defining stress as a prelude to mapping its neurocircuitry: no help from allostasis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **29**, 195–200.

de Kloet, E. R., Joëls, M. and Holsboer, F. (2005). Stress and the brain: from adaptation to disease. *Nature Reviews Neuroscience* **6**, 463–475.

de Kloet, E. R., Vreugdenhil, E., Oitzl, M. S., et al. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* **19**, 269–301.

Joëls, M., Pu, Z., Wiegert, O., et al. (2006). Learning under stress: how does it work? *Trends in Cognitive Science* **10**, 152–158.

Karst, H., Berger, S., Turiault, M., et al. (2005). Mineralocorticoid receptors are indispensable for nongenomic modulation of hippocampal glutamate transmission by corticosterone. *Proceedings of the National Academy of Sciences USA* **102**, 19204–19207.

Kruk, M. R., Halász, J., Meelis, W., et al. (2004). Fast positive feedback between the adrenocortical stress response and a brain mechanism involved in aggressive behaviour. *Behavioral Neuroscience* **118**, 1062–1070.

McEwen, B. S. (2004). Protection and damage from acute and chronic stress: allostasis and allostatic overload and relevance to the pathophysiology of psychiatric disorders. *Annals of the New York Academy of Sciences* **1032**, 1–7.

McEwen, B. S. and Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior* **43**, 355–367.

Sapolsky, R. M., Romero, L. M. and Munck, A. U. (2000). How do glucocorticoids influence stress responses? integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews* **21**, 55–89.

Steptoe, A. and Wardle, J. (2005). Positive affect and biological function in everyday life. *Neurobiology of Aging* **26**, 108–112.

VanPraag, H. M., de Kloet, E. R. and Van Os, J. (2004). *Stress, the brain and depression.* Cambridge, UK: Cambridge University Press.

Stress, Beneficial Effects of

S Joseph
University of Nottingham, Nottingham, UK
P A Linley
University of Leicester, Leicester UK

© 2007 Elsevier Inc. All rights reserved.

Types of Benefit
Prevalence
Etiological Factors
Theoretical Developments
Clinical Applications

Glossary

Adversarial growth	Increased psychological functioning and personal development following exposure to a stressful and traumatic event; also called posttraumatic growth or stress-related growth.
Assumptive world	Cognitive schemata about the self and world that determine how meaning is perceived.

A number of literatures, religions, and philosophies throughout human history have conveyed the idea that there are benefits to be found in suffering. However, it is only relatively recently that the topic of benefit finding following stress and trauma has become a research focus. Initially there was some debate over the validity of the concept of benefit finding, with some commentators questioning whether it was anything more than positive illusion and self-deception. But there is now convincing evidence that people often experience benefits following stress and trauma. These benefits have been variously labeled adversarial growth, benefit finding, flourishing, heightened existential awareness, perceived benefits, positive by-products, positive changes, positive meaning, posttraumatic growth, quantum change, self-renewal,

stress-related growth, thriving, and transformational coping. These terms have been used interchangeably to refer not just to recovery following stressful and traumatic events but to how events can sometimes serve as a springboard to a higher level of psychological functioning.

Types of Benefit

Three broad interrelated dimensions of benefit finding have been discussed. First, relationships are enhanced in some way; for example, people now value their friends and family more and feel an increased compassion and altruism toward others. Second, people change their views of themselves in some way; for example, they have a greater sense of personal resiliency, wisdom, and strength, perhaps coupled with a greater acceptance of their vulnerabilities and limitations. Third, people report changes in life philosophy; for example, they find a fresh appreciation for each new day and renegotiate what really matters in the full realization that their life is finite, including possibly changing their spiritual beliefs. Various psychometric self-report instruments have been developed with which to assess positive changes and personal growth following adversity: the Changes in Outlook Questionnaire, the Perceived Benefit Scale, the Posttraumatic Growth Inventory, the Stress-Related Growth Scale, and the Thriving Scale. The development of these instruments has led to further empirical investigation.

Prevalence

Prevalence varies among studies due to samples, contexts, and methodological issues, but it is expected that anywhere between 30 and 70% of people may experience some form of benefit following traumatic events, although the benefits are not usually reported in the immediate aftermath but at some later time. Also, the benefits are likely to coexist with psychological distress; the experience of benefits is not mutually exclusive with other more negative psychological consequences.

Etiological Factors

Studies have reported benefits following a range of stressful and traumatic events, for example, bereavements; accidents and disasters; chronic and life-threatening illnesses; sexual, physical, and emotional abuse in childhood; sexual assaults; and wars and conflicts. Research has also begun to document the correlates and predictors of growth, with evidence pointing to the importance of stress-appraisal, coping, and personality variables, with the more optimistic, self-efficacious, extraverted, and stable people, who use more spiritual, accepting, emotionally expressive, and emotionally focused forms of coping, being more likely to report benefits. Social support is thought to be important too.

Theoretical Developments

Tedeschi and Calhoun's functional-descriptive model shows how traumatic events serve as seismic challenges to the pretrauma schema by shattering prior goals, beliefs, and ways of managing emotional distress. When these schemas are shattered in this way, this leads to ruminative activity as people try to make sense of what has happened and to deal with their emotional reactions to the trauma. In the initial stages, this ruminative activity is more automatic than deliberate. Although distressing, automatic ruminative activity is indicative of cognitive activity that is directed at rebuilding the pretrauma schema. This ruminative process is influenced by social support networks, which provide sources of comfort and relief, as well as being influenced by new coping behaviors and available options for the construction of new posttrauma schemas. Successful coping at this stage facilitates disengagement from goals that are now unreachable and from beliefs that are no longer tenable in the posttrauma environment, together with decreased emotional distress. As successful coping aids adaptation, the initial ruminative activity that was characterized by its automatic nature shifts toward a more effortful ruminative activity. This effortful ruminative activity is characterized by narrative development, part of which may be the search for meaning.

Introducing a metatheoretical perspective, Joseph and Linley proposed that human beings are active, growth-oriented organisms. The theory proposes that individuals are intrinsically motivated to cognitively accommodate their psychological experiences and thus find benefit. The confrontation with a stressful and traumatic event has a shattering effect on the person's assumptive world. When the social environment is able to provide for the basic human needs of autonomy, competence, and relatedness, then models of the self and world transform to accommodate the new trauma-related information. When the social environment does not provide support for these basic human needs, the process of accommodation is thwarted and benefits are less likely to be experienced.

Similarly, but from a biopsychosocial-evolutionary perspective, Christopher regarded growth, rather than pathology, as the normal outcome of the traumatic stress response. Reviewing and synthesizing the trauma literature, Christopher articulated seven interconnected theoretical conclusions:

1. Stress is best understood as a prerational form of biopsychological feedback.
2. The normal outcome of traumatic stress is growth.
3. Any resultant psychopathology is a function of a maladaptive modulation of the stress response.
4. Trauma transforms individuals on the biological as well as psychological levels.
5. Biological processes underlying stress response are universal, but there are specific dynamics that are a function of the individual.
6. Biological changes can occur even if there are no psychological changes.
7. Rationality is humanity's evolutionary newest and most sophisticated stress-reduction mechanism.

Clinical Applications

The interest in benefit finding can be seen as part of the wider positive psychology movement. Positive psychology is seen as a balance to mainstream psychology, which has been for too long overly concerned with the negative aspects of human experience. Research in positive psychology is now beginning to focus on its applications, and the study of benefit finding promises to have important applications for clinical, counseling, and health psychologists, who when working with clients who have experienced stressful and traumatic events need not be concerned only with helping to alleviate distress but also with helping to facilitate more positive functioning.

The importance of this avenue of investigation is seen in the seminal work of Affleck and colleagues, who found that perceived benefits at 7 weeks following a heart attack significantly predicted less chance of heart attack recurrence and lower general health morbidity at an 8-year follow-up. Experimental studies to test whether the principles of growth might somehow be introduced as part of a clinical intervention are encouraging. For example, Stanton and colleagues randomly assigned breast cancer patients to one of two groups, one group to write about the facts of the cancer experience and the other to write about positive thoughts and feelings regarding the experience. It was found that those assigned to write about their positive experiences had significantly fewer medical appointments for cancer-related morbidities 3 months later.

Clinicians should be aware of the potential for positive change in their clients following stress and trauma. But it must also be recognized that adversity does not lead to positive change for everyone. Therefore, therapists need to be careful not to inadvertently imply that patients have in some way failed by not making more of their experience or that there is anything inherently positive in their experiences. Personal growth after trauma should be viewed as originating not from the event, but from within the patients themselves through the process of their struggle with the event and its aftermath.

See Also the Following Articles

Bereavement; Caregivers, Stress and; Cognitive Behavioral Therapy; Integrative Medicine (Complementary and Alternative Medicine); Health Behavior and Stress; Mental Stress Testing; Religion and Stress; Self-Esteem, Stress and Emotion; Recovery from Stress; Stress of Self Esteem.

Further Reading

Affleck, G., Tennen, H., Croog, S., et al. (1987). Causal attribution, perceived benefits, and morbidity after a heart attack: an 8-year study. *Journal of Consulting and Clinical Psychology* **55**, 29–35.

Calhoun, L. G. and Tedeschi, R. G. (1999). *Facilitating posttraumatic growth: a clinician's guide*. Mahwah, NJ: Lawrence Erlbaum.

Christopher, M. (2004). A broader view of trauma: a biopsychosocial-evolutionary view of the role of the traumatic stress response in the emergence of pathology and/or growth. *Clinical Psychology Review* **24**, 75–98.

Cordova, M. J., Cunningham, L. L. C., Carlson, C. R., et al. (2001). Posttraumatic growth following breast cancer: a controlled comparison study. *Health Psychology* **20**, 176–185.

Frazier, P., Tashiro, T., Berman, M., et al. (2004). Correlates of levels and patterns of positive life changes following sexual assault. *Journal of Consulting and Clinical Psychology* **72**, 19–30.

Joseph, S. and Linley, P. A. (2005). Positive adjustment to threatening events: an organismic valuing theory of growth through adversity. *Review of General Psychology* **9**, 262–280.

Linley, P. A. and Joseph, S. (2004). Positive change following trauma and adversity: a review. *Journal of Traumatic Stress* **17**, 11–21.

O'Leary, V. E., Alday, C. S. and Ickovics, J. R. (1998). Models of life change and posttraumatic growth. In: Tedeschi, R. G., Park, C. L. & Calhoun, L. G. (eds.) *Posttraumatic growth: positive changes in the aftermath of crisis*, pp. 127–151. Mahwah, NJ: Lawrence Erlbaum.

Park, C. L., Cohen, L. H. and Murch, R. (1996). Assessment and prediction of stress-related growth. *Journal of Personality* **64**, 71–105.

Seligman, M. E. P. and Csikszentmihalyi, M. (2000). Positive psychology: an introduction. *American Psychologist* **55**, 5–14.

Stanton, A. L., Danoff-Burg, S., Sworowski, L. A., et al. (2002). Randomized controlled trial of written emotional expression and benefit finding in breast cancer patients. *Journal of Clinical Oncology* **20**, 4160–4168.

Tedeschi, R. G. and Calhoun, L. G. (2004). A clinical approach to posttraumatic growth. In: Linley, P. A. & Joseph, S. (eds.) *Positive psychology in practice*, pp. 405–419. Hoboken, NJ: John Wiley.

Tedeschi, R. G. and Calhoun, L. G. (2004). Posttraumatic growth: conceptual foundations and empirical evidence. *Psychological Inquiry* **15**, 1–18.

Tedeschi, R. G., Park, C. L. and Calhoun, L. G. (eds.) (1998). *Posttraumatic growth: positive changes in the aftermath of crisis*. Mahwah, NJ: Lawrence Erlbaum.

Tennen, H. and Affleck, G. (2002). Benefit-finding and benefit-reminding. In: Snyder, C. R. & Lopez, S. J. (eds.) *Handbook of positive psychology*, pp. 584–597. New York: Oxford University Press.

Stress, Definitions and Concepts of

B S McEwen
Rockefeller University, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by B S McEwen, volume 3, pp 508–509, © 2000, Elsevier Inc.

Stress involves a stressor and a stress response. A stressor may be a physical insult, such as trauma or injury, or physical exertion, particularly when the body is being forced to operate beyond its capacity. Other physical stressors include noise, overcrowding, and excessive heat or cold. Stressors also include primarily psychological experiences such as time-pressured tasks, interpersonal conflict, unexpected events, frustration, isolation, and traumatic life events, and all of these types of stressors may produce behavioral responses and evoke physiological consequences such as increased blood pressure, elevated heart rate, increased cortisol levels, impaired cognitive function, and altered metabolism.

Behavioral responses to stressors may decrease risk and get the individual out of trouble or involve health-promoting activities such as a good diet and regular exercise, but they may also include responses that exacerbate the physiological consequences of stress such as self-damaging behaviors (e.g., smoking, drinking, overeating, or consuming a rich diet) and risk-taking behaviors (e.g., driving an automobile recklessly). The physiological stress responses include primarily the activation of the autonomic nervous system and the hypothalamic-pituitary-adrenal (HPA) axis, leading to increased blood and tissue levels of catecholamines and glucocorticoids. It is these physiological responses that have both protective and damaging effects (see **Allostasis and Allostatic Load**).

There are two important features of the physiological stress response. The first is turning it on in amounts that are adequate to the challenge (see **Allostasis and Allostatic Load**). The second is turning off the response when it is no longer needed. The physiological mediators of the stress response, namely, the catecholamines of the sympathetic nervous system and the glucocorticoids from the adrenal cortex, initiate cellular events that promote adaptive changes in cells and tissues throughout the body, which in turn protect the organism and promote survival. However, too much stress or inefficient operation of the acute responses to stress can cause wear and tear and exacerbate disease processes.

There are enormous individual differences in interpreting and responding to what is stressful, as well as individual differences in the susceptibility to diseases, in which stress may play a role. Genetic predispositions exist that increase the risk of certain disorders. In addition, developmental process, such as prenatal stress or nurturing postnatal experiences, contribute to the life-long responsiveness of the behavioral and physiological responses to stressors. Furthermore, experiences throughout the life course that result in memories of particularly unpleasant or pleasant situations combine with the genetic and developmental influences to produce large differences among individuals in how they react to stress and what the long-term consequences may be.

See Also the Following Articles

Allostasis and Allostatic Load; Homeostasis.

Further Reading

Mason, J. (1959). Psychological influences on the pituitary-adrenal cortical system. In: Pincus, G. (ed.) *Recent progress in hormone research*, pp. 345–389. New York: Academic Press.

Selye, H. (1936). A syndrome produced by diverse nocuous agents. *Nature* **138**, 32.

Selye, H. (1955). Stress and disease. *Science* **122**, 625–631.

Stress, General Overview *See:* Stress, Definitions and Concepts of; Stress Effects, Overview.

Stress, Insulin Resistance, and Type II Diabetes

C Tsigos and I Kyrou
Evgenidion Hospital, Athens University Medical School, Athens, Greece

© 2007 Published by Elsevier Inc.

The Stress Response
Depression and Diabetes
Stressful Life Events and Diabetes
Inflammation and Diabetes
Conclusion

Glossary

Adipokines	Cytokine-like molecules produced and secreted by the adipose tissue (e.g., leptin, adiponectin, resistin, tumor necrosis factor α [TNF-α], interleukin-6 [IL-6]).
Insulin resistance	Decreased ability of insulin to act on its target tissues, resulting in compensatory hyperinsulinemia. Considered to be the earlier abnormality in the natural history of type II diabetes.
Metabolic syndrome	Syndrome characterized by the combination of central obesity, insulin resistance, dyslipidemia, and hypertension and correlated with high mortality rates due to cardiovascular complications, which aptly justify the description of the syndrome as the deadly quartet.
Visceral obesity	Obesity characterized by fat accumulation predominantly in the abdominal area, also called central or abdominal obesity.

The Stress Response

Stress is defined as a state of disharmony or threatened homeostasis. The concepts of stress and homeostasis can be traced back to ancient Greek history; however, the integration of these notions with related physiological and pathophysiological mechanisms and their association with specific illnesses are much more recent.

When faced with excessive stress, whether biological, physical, or psychological, a subject's adaptive responses attain a stereotypic nonspecific nature, referred to by Selye as the general adaptation syndrome. During stress, attention is enhanced and the brain focuses on the perceived threat. Cardiac output and respiration are accelerated, catabolism is increased, and blood flow is redirected to provide the highest perfusion and fuel to the aroused brain, heart, and muscles.

The central control stations of the stress system are located in the hypothalamus and the brain stem and include the parvocellular corticotropin-releasing hormone (CRH) and arginine-vasopressin (AVP) neurons of the paraventricular nuclei of the hypothalamus and the locus ceruleus-norepinephrine system (central sympathetic system) (**Figure 1**). The hypothalamic-pituitary-adrenal (HPA) axis, together with the efferent sympathetic/adrenomedulary system, represent the effector limbs, via which the brain influences all body organs during exposure to threatening stimuli.

The Hypothalamic-Pituitary-Adrenal (HPA) Axis

The hypothalamus controls the secretion of adrenocorticotropic hormone (ACTH) from the anterior pituitary, which, in turn, stimulates the secretion by the adrenal cortex of glucocorticoids, mainly cortisol in humans. The principal hypothalamic stimulus to the pituitary-adrenal axis is CRH. AVP is a potent synergistic factor with CRH in stimulating ACTH secretion.

In nonstressful situations, both CRH and AVP are secreted into the hypophysial portal system in a circadian, pulsatile fashion, with a frequency of about two to three secretory episodes per hour. Under resting conditions, the amplitude of the CRH and AVP pulses increases in the early morning hours, resulting in ACTH and cortisol secretory bursts in the general circulation. These diurnal variations are perturbed by changes in lighting, feeding schedules, and activity and are disrupted by stress.

Circulating ACTH is the key regulator of glucocorticoid secretion by the adrenal cortex. Glucocorticoids are the final effectors of the HPA axis and participate in the control of whole body homeostasis and the organism's response to stress. They play a key regulatory role in the basal activity of the HPA axis

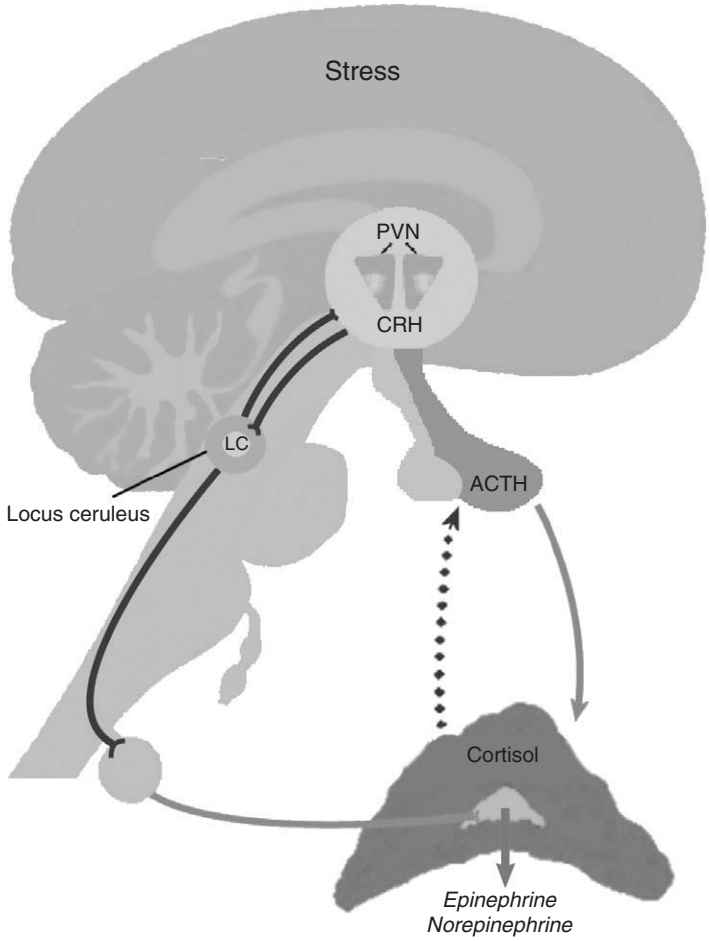


Figure 1 Schematic representation of the stress system, highlighting the major central and peripheral components involved in the stress response. PVN, paraventricular nucleus; CRH, corticotropin-releasing hormone; LC, locus ceruleus/norepinephrine-sympathetic system; ACTH, adrenocorticotropic hormone. Activation is represented by solid lines and inhibition by dashed lines.

and in the termination of the stress response by acting at the hypothalamus and the pituitary gland and at higher brain centers. Glucocorticoids exert these actions via their glucocorticoid receptors and the rigorous regulation of a plethora of genes that are directly or indirectly implicated in several metabolic pathways, either activating them through positive glucocorticoid responsive elements (GREs) (e.g., genes encoding enzymes of glucose, lipid and amino acid metabolism, lipoprotein receptors, leptin) or repressing them through negative GREs (e.g., CRH, proopiomelanocortin, various cytokines).

The Sympathetic Nervous System

The autonomic nervous system provides a rapidly responding mechanism to control a wide range of functions. Cardiovascular, respiratory, gastrointestinal, renal, endocrine, and other systems are regulated by the sympathetic nervous system, the parasympathetic system, or both. Sympathetic innervation of peripheral organs is derived from the efferent preganglionic fibers, whose cell bodies lie in the intermediolateral column of the spinal cord. These nerves synapse in the bilateral chains of sympathetic ganglia, with postganglionic sympathetic neurons that richly innervate the smooth muscle of the vasculature, the heart, skeletal muscles, kidney, gut, fat, and many other organs. The sympathetic system also has a humoral contribution, providing most of the circulating epinephrine and some of the norepinephrine from the adrenal medulla (Figure 1).

Consequences of Chronic Stress

Generally, the stress response with the resultant activation of the HPA and the sympathetic nervous system axes is meant to be acute or at least of a limited duration. The time-limited nature of this process renders its accompanying antireproductive, antigrowth, catabolic, and immunosuppressive effects temporarily beneficial rather than damaging. In contrast, chronicity of stress system activation would lead to the syndromal state that Selye described in 1936.

The prototypic example of chronic hyperactivation of the stress system is manifested in melancholic depression. Owing to a chronically hyperactive stress system, patients with melancholic depression may sustain several severe somatic sequelae, such as osteoporosis, features of the metabolic syndrome (e.g., visceral fat accumulation, dyslipidemia, increased blood pressure, impaired glucose tolerance), varying degrees of atherosclerosis, immunosuppression, and certain infectious and neoplastic diseases. When not treated, these patients have a compromised life expectancy curtailed by 15–20 years after excluding suicides.

In addition to melancholic depression, a spectrum of other conditions may be associated with increased and prolonged activation of the HPA axis, including childhood sexual abuse, anorexia nervosa with or without malnutrition, obsessive-compulsive disorder, panic anxiety, chronic active alcoholism, and alcohol and narcotic withdrawal.

More specifically, in the context of a mobilized adaptive stress response, glucocorticoids exert a broad spectrum of primarily catabolic effects as part of a generalized effort to utilize every available energy resource against the challenge enforced by the stressors. Thus, glucocorticoids increase hepatic gluconeogenesis and elevate plasma glucose concentrations, induce lipolysis (although they favor abdominal and dorsocervical fat accumulation), and cause protein degradation at multiple tissues (e.g., muscle, bone, skin) to provide amino acids that would be used as an additional substrate for oxidative pathways. In addition to their direct catabolic actions, glucocorticoids also antagonize the beneficial anabolic actions of growth and thyroid hormones, insulin, and sex steroids on their target tissues (Figure 2). This stress-related shift of the metabolism toward a catabolic state normally reverses upon retraction of the posed stressor. Chronic exposure to stress and HPA axis activation, however, is expected to progressively

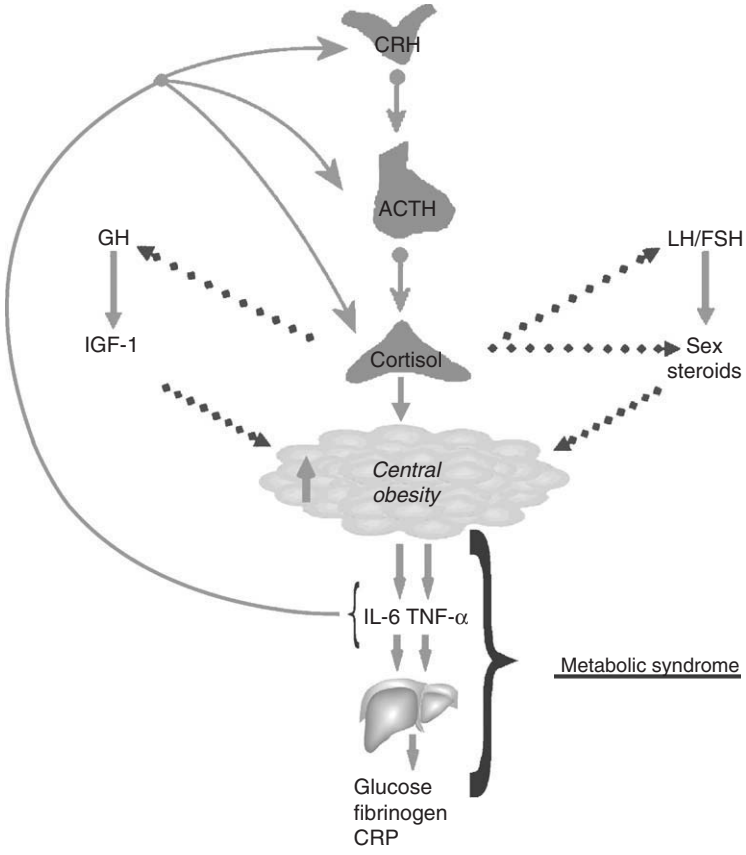


Figure 2 Schematic representation of the detrimental effects of chronic stress on adipose tissue that may lead to clinical manifestations of the metabolic syndrome. Chronic HPA axis activation favors visceral obesity, with cortisol suppressing the beneficial effects of both the gonadal and growth hormone axis and directly stimulating adipocyte proliferation. Hypersecretion of TNF-α and IL-6 is proportional to the cytokines further stimulate the HPA axis activation, thus forming a deleterious vicious cycle. CRH, corticotropin-releasing hormone; ACTH, adrenocorticotropic hormone; TNF-α, tumor necrosis factor α; IL-6, interleukin-6; CRP, C-reactive protein. Stimulatory effects are represented by solid arrows and inhibitory effects by dashed arrows.

increase visceral adiposity, decrease lean body (muscle and bone) mass, and cause insulin resistance (Figure 2). Interestingly, the phenotype of Cushing's syndrome, characterized by abdominal and trunk fat accumulation and decreased lean body mass, in combination with various other manifestations of the metabolic syndrome (insulin resistance, dyslipidemia, hypertension, hypercoagulability, and hypercytokinemia), is present in a variety of pathophysiological conditions, categorized as pseudo-Cushing's states. This cushingoid phenotype, under certain conditions of chronic stress, could be attributed to stress-induced mild hypercortisolism, possibly combined with increased peripheral tissue sensitivity to glucocorticoids, the latter depending on the tissue-specific concentration and activity of glucocorticoid receptors, cortisol-binding protein (CBG), and 11β-hydroxysteroid dehydrogenase type 1 (11β-HSD1). Interestingly, 11β-HSD1 is an isoform expressed primarily in the liver, adipose tissue, kidneys, and brain that catalyzes the conversion of the inactive cortisone to active cortisol in humans and, thus, locally augments the concentration of the bioactive hormone. This tissue-specific amplification of the cortisol action, mediated by 11β-HSD1, appears to be very important for determining the tissue sensitivity to glucocorticoids and hence to the HPA axis hyperactivation, especially under conditions that potentiate 11β-HSD1 activity, such as visceral obesity.

Depression and Diabetes

Until very recently, the attention of the scientific community had been focused on the emotional consequences of diabetes mellitus, rather than on the emotional causes of the disease. Depression rates are roughly twice as high among persons with diabetes as among persons without diabetes and are higher than those for most other chronic physical illnesses. A meta-analysis of 42 eligible studies showed that the presence of diabetes doubles the odds of comorbid depression, a stable finding in both controlled and uncontrolled studies not affected by sex, subject source, assessment method, or type of diabetes. Yet most of these findings are based on cross-sectional data, so they actually tell us very little about causal relationships.

Recently, however, significant evidence from population-based prospective studies in the United States and Japan shows that being depressed at baseline doubles the risk of subsequently developing type II diabetes and that this association is not explained by other risk factors for type II diabetes, including obesity, activity level, smoking, alcohol consumption, and family history of type II diabetes. Thus, in some circumstances, depression may precede the onset of type II diabetes and may play an important role in its development. As described previously, depressive disorders are accompanied by increased sympathetic and sympatho-adrenal system activity, which is known to be associated with increased blood glucose and impaired glucose tolerance. Depressive disorders have also been associated with increased activity of the HPA axis, which results in an increased release of glucocorticoids, decreased glucose uptake, and elevated glucose levels. Depression may impair the ability to handle a carbohydrate load and then increase the risk of developing type II diabetes through a greater release of the counterregulatory hormones. More importantly, depression is associated with visceral fat accumulation, which is considered the key abnormality underlying the development of insulin resistance and the increased risk for type II diabetes and atherosclerosis. Changes in diet (e.g., unhealthy eating, alcohol consumption) and physical activity associated with chronic depression may also contribute to the glucose abnormalities that may develop.

Interestingly, effective treatment of depression in diabetic patients appears to produce a better glycemic control that is not due to differential effects of treatment on weight. Moreover, in controlled trials of cognitive-behavioral therapy, related improvements in depression produced significant improvements in glycemic control.

Stressful Life Events and Diabetes

Several studies have demonstrated a relationship of stress to glycemic control in samples of patients with type II diabetes. Stress can be managed through the use of behavioral stress management programs or through the administration of anxiolytic medications, both of which have been reported to improve long-term glycemic control in patients with type II diabetes. It is of note that staff-intensive individual training is not always required for a successful behavioral intervention, as group programs in a real-world setting can also result in clinically significant benefits for patients with type II diabetes.

A large recent study (the Hoorn Study) showed that a high number of rather common major life events, consistent with chronic psychological stress, during the past 5 years was indeed related to a higher prevalence of previously unknown type II diabetes, as well as to increased levels of visceral fat accumulation. These associations were independent of family history of diabetes, physical activity, heavy alcohol use, or low level of education. Furthermore, behavioral stress management programs, whether individualized or in groups, usually help.

Inflammation and Diabetes

It has been known for several decades that inflammatory stress is associated with concurrent activation of the HPA axis. Indeed, cytokines and other humoral mediators of inflammation are potent activators of the central stress response, constituting the afferent limb of the feedback loop through which the immune/inflammatory system and the central nervous system communicate (Figure 3).

All three inflammatory cytokines tumor necrosis factor α (TNF-α), interleukin-1β, and interleukin-6 (IL-6) can cause stimulation of the HPA axis alone, or in synergy with each other. There is evidence to suggest that IL-6, the main endocrine cytokine, plays the major role in the immune stimulation of the axis. Some of the activating effects of cytokines on the HPA axis may be exerted indirectly by stimulation of the central catecholaminergic pathways. Also, activation of peripheral nociceptive, somatosensory, and visceral afferent fibers would lead to stimulation of both the catecholaminergic and CRH neuronal systems via ascending spinal pathways.

In addition, TNF-α interferes directly with insulin signaling causing insulin resistance, and, together with the other inflammatory cytokines, has direct actions on the endothelium promoting the atherogenic process. Interestingly, TNF-α is overproduced by adipose tissue in obesity and appears to be one of the molecular links between obesity insulin resistance/diabetes and atherosclerosis. Increased production of other adipokines, such as resistin, visfatin, and leptin, and decreased production of adiponectin with expanding obesity, further increase the inflammatory load and the risk for insulin resistance and atherosclerosis.

Indeed, an increasing body of evidence supports a possible role for inflammation in the pathogenesis of type II diabetes. Thus, in the Women's Health Study, an ongoing U.S. primary prevention randomized clinical trial initiated in 1992, elevated levels of C-reactive protein (CRP) and IL-6 predicted the development of type II diabetes. Other common markers of inflammation (fibrinogen, haptoglobin, white-cell count, sialic acid) are also associated prospectively with the development of diabetes in adults. Furthermore, chronic subclinical inflammation relates independently to insulin resistance, even in nondiabetic subjects. It is therefore likely that anti-inflammatory therapy may be beneficial in improving insulin resistance and preventing the development of type II diabetes. Interestingly, thiazolidinediones, a new class of insulin sensitizers, also have anti-inflammatory properties, reducing CRP and TNF-α levels. Moreover, the observed significant reductions in the hazard of becoming diabetic with statins may be attributed at least partly to the anti-inflammatory properties of these drugs.

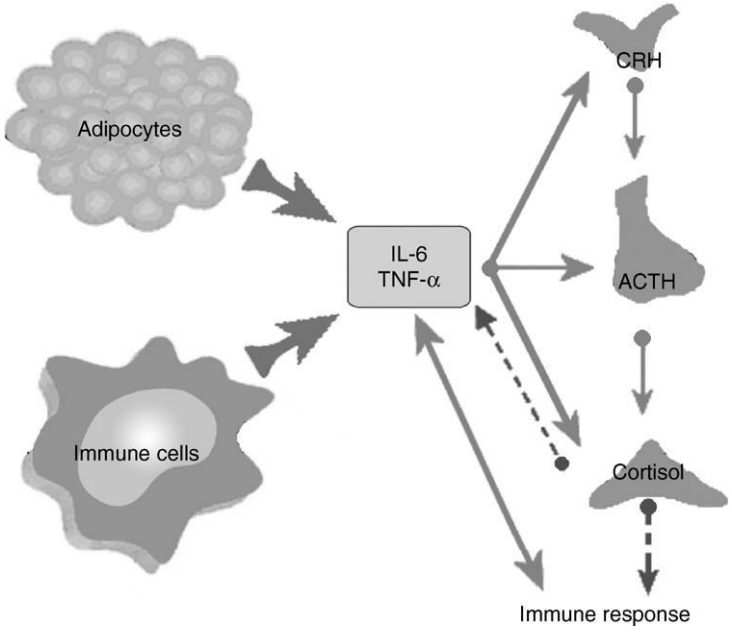


Figure 3 Schematic representation of the interactions between the pro-inflammatory cytokines, the HPA axis, and the immune system response. TNF-α and IL-6 secretion by the adipocytes is continuous and proportional to body mass index (BMI). TNF-α and IL-6 secretion by the immune cells at inflammatory sites occurs in a cascade-like fashion, with TNF-α appearing first followed by IL-6. TNF-α, tumor necrosis factor α; IL-6, interleukin-6; CRH, corticotropin-releasing hormone; ACTH, adrenocorticotropic hormone. Stimulatory effects are represented by solid arrows and inhibitory effects by dashed arrows.

Conclusion

Chronic stress seems to contribute to the development of visceral obesity and, hence, of type II diabetes and atherosclerosis, and, conversely, the accumulation of visceral adipose tissue constitutes a chronic stressful state. The mechanisms that mediate this documented reciprocal relation and cause the derangement of the metabolic homeostasis are the subject of many recent research efforts, and their characterization will hopefully provide novel insights into the pathogenesis of type II diabetes and cardiometabolic complications and help the quest for more specific and effective therapeutic interventions. Moreover, treating depression in and providing effective stress management programs for people predisposed to diabetes may add a new dimension of primary prevention of type II diabetes to the treatment of traditional risk factors such as obesity, diet, and physical inactivity.

See Also the Following Articles

11β-Hydroxysteroid Dehydrogenases; Atherosclerosis; Metabolic Syndrome; Metabolic Syndrome and Stress; Inflammation.

Further Reading

Anderson, R. J., Freedland, K. E., Clouse, R. E. and Lustman, P. J. (2001). The prevalence of comorbid depression in adults with diabetes. A meta-analysis. *Diabetes Care* **24**, 1069–1078.

Barzilay, J. I., Abraham, L., Heckbert, S. R., et al. (2001). The relation of markers of inflammation to the development of glucose disorders in the elderly. The Cardiovascular Health Study. *Diabetes* **50**, 2384–2389.

Bjorntorp, P., Holm, G. and Rosmond, R. (1999). Hypothalamic arousal, insulin resistance and type 2 diabetes mellitus. *Diabetic Medicine* **16**, 373–383.

Buren, J. and Eriksson, J. W. (2005). Is insulin resistance caused by defects in insulin's target cells or by a stressed mind? *Diabetes/Metabolism Research and Reviews* **21**, 487–494.

Chrousos, G. P. (1995). The hypothalamic-pituitary-adrenal axis and immune-mediated inflammation. *New England Journal of Medicine* **332**, 1351–1362.

Chrousos, G. P. (2000). The role of stress and the hypothalamic-pituitary-adrenal axis in the pathogenesis of the metabolic syndrome: neuro-endocrine and target tissue-related causes. *International Journal of Obesity* **24**(Supplement 2), S50–*55*.

Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress system disorders: overview of behavioral and physical homeostasis. *Journal of the American Medical Association* **267**, 1244–1252.

Moller, D. E. and Kaufman, K. D. (2005). Metabolic syndrome: a clinical and molecular perspective. *Annual Reviews in Medicine* **56**, 45–62.

Rubin, R. R. and Peyrot, M. (2002). Was Willis right? Thoughts on the interaction of depression and diabetes. *Diabetes/Metabolism Research and Reviews* **18**, 173–175.

Schmidt, M. I., Duncan, B. B., Sharrett, A. R., et al. for the ARIC Investigators (1999). Markers of inflammation and prediction of diabetes mellitus in adults (Atherosclerosis Risk in Communities study): a cohort study. *Lancet* **353**, 1649–1652.

Seckl, J. R., Morton, N. M. and Chapman, K. E. (2004). Glucocorticoids and 11beta-hydroxysteroid dehydrogenase in adipose tissue. *Recent Progress in Hormone Research* **59**, 359–393.

Surwit, R. S., van Tilburg, M. A. L., Zucker, N., et al. (2002). Stress management improves long-term glycemic control in type 2 diabetes. *Diabetes Care* **25**, 30–34.

Tsigos, C. and Chrousos, G. P. (2002). Hypothalamic-pituitary-adrenal axis, neuroendocrine factors and stress. *Journal of Psychosomatic Research* **53**, 865–871.

Tsigos, C. and Kyrou, I. (2006). Hypothalamic-pituitary-adrenal axis, cytokines and the metabolic syndrome. *Obesity and Metabolism* **2**, 116–126.

Tsigos, C., Kyrou, I. and Chrousos, G. P. (2005). Stress, endocrine manifestations, and diseases. In: Cooper, C. L. (ed.) *Handbook of stress, medicine, and health* (2nd edn., pp. 101–131). Boca Raton, FL: CRC Press.

Stress, NPY, and Cardiovascular Diseases

Z Zukowska
Georgetown University Medical School, Washington, DC, USA

© 2007 Elsevier Inc. All rights reserved.

Structure and Localization of NPY in the Body
The NPY Receptor System and Major Actions
Sources of NPY Release in the Circulation
Stress-Induced Release of NPY from Peripheral Sympathetic Nerves
Chronic Stress, NPY, and Vascular Remodeling after Balloon Angioplasty
Critical Role of Y1 Receptors in Immune System and Angioplasty-Induced Restenosis
Chronic Stress and Angiogenesis: Implications for Ischemic Revascularization and Adipogenesis
Summary and Conclusion

Glossary

Adipogenesis	The growth of adipose tissue, increasing fat mass.
Angiogenesis	The formation of new vessels from preexisting ones; called ischemic angiogenesis in areas of poor blood flow (ischemia).
Atherosclerosis	The inflammatory response of blood vessels; the buildup of atherosclerotic plaques causing a lesion that contains the abnormal growth of vascular smooth muscle cells (neointima), deposition of matrix proteins, aggregation of platelets (thrombus formation), and formation of lipids.
Balloon angioplasty	A procedure used by cardiologists that involves the insertion of a catheter whose balloon can be inflated to re-open an artery that was closed (occluded) by an atherosclerotic plaque.
Macrophages	Immune cells with phagocytotic capabilities.
Neuropeptide Y (NPY)	A 36-amino-acid peptide present primarily in the neurons of the brain and in the peripheral nervous system, rich in tyrosine (Y, in the one-letter code used for amino acids).
Obesity	Increased fat mass and weight above 30% of normal.
Restenosis	The repeated closure of a vessel previously opened by balloon angioplasty or by other vascular procedures such as the insertion of a stent.
Stress	The real or perceived alteration of internal or external environment that elicits bodily responses (stress responses).
Sympathetic nerves	Part of the autonomic or vegetative nervous system that innervates most blood vessels and organs of the body; essential in body's responses to stress.
Vascular remodeling	Any change to vessel structure such as the excessive or abnormal growth of vascular smooth muscle in the medial layer of the vessel or the sprouting of new vessels (angiogenesis).

Structure and Localization of NPY in the Body

Neuropeptide Y (NPY) is a 36-amino-acid brain peptide, one of the most abundant in the body. It is formed from a 97-amino-acid peptide, preproNPY, which contains a signal peptide targeting the precursor molecule to the endoplasmic reticulum. Specific convertases subsequently cleave the preproNPY to form a mature and amidated NPY(1–36), which is biologically active. An additional processing product is the C-terminal peptide of NPY (CPON), a peptide whose actions are relatively unknown. Inside neurons, mature NPY(1–36) is stored in large dense core vesicles and emptied into the synaptic cleft during nerve depolarization. Recent research shows that, in addition to its neuronal sources, NPY is expressed extraneuronally in the endothelial cells and platelets of some species, where it acts as an autocrine and paracrine regulator of cell functions.

In the central nervous system, NPY is richly expressed in the hypothalamus, hippocampus, brain stem, and cortex, and it is colocalized with classical neurotransmitters such as catecholamines and GABA. In the periphery, NPY is present in all the sympathetic nerves innervating the blood vessels and in the chromaffin cells of the adrenal medulla, where it is colocalized with norepinephrine (NE) and epinephrine, respectively. The most abundant vascular NPY-ergic innervation is found in small arteries in the kidneys and the mesenteric bed. The third, least source of NPY in the body is the gut parasympathetic and enteric nerves, in which NPY is present in small and large intestines. In the gastrointestinal tract, two other peptides belong to the same family as NPY: peptide YY (PYY) and pancreatic polypeptide (PP). They act as hormones rather than as neurotransmitters; hence, they are not a major focus of this review.

The whole NPY peptide family evolved from an ancient ancestor gene and has maintained a high degree of conservation of structure. NPY, in particular, is 90–99% homologous among a wide range of species; for example, rat and human NPYs are identical. The critical C-terminal portion of NPY, as well as PYY, which shares most of its affinity for its receptors, has been 100% conserved throughout evolution, suggesting that it plays an important physiological role.

The NPY Receptor System and Major Actions

Determining what NPY's physiological role is has been a challenge. Early Swedish studies suggested potent cardiovascular effects for NPY – arterioconstriction and increased blood pressure, which are resistant to the blockade of α-adrenergic receptors. In a landmark experiment in 1982 on the cat submandibular gland, Lundberg et al. showed that the effect of NPY is mimicked by prolonged sympathetic nerve stimulation in the presence of total adrenergic blockade and suggested that NPY is a mediator of nonadrenergic sympathetic transmission. Determining the significance of these actions in the cardiovascular system required the cloning of the NPY receptors and the development of selective receptor inhibitors.

The first receptor that was cloned for rat, mouse, and humans was the Y1 receptor. It is a Gi-coupled receptor that is predominant in the blood vessels, where it mediates vasoconstriction and the potentiation of vasoconstriction exerted by NE and other agents. The Y1 receptor is also widely expressed in neurons in the central and peripheral nervous systems, where it mediates a range of functions from anxiolysis and appetite stimulation to neuronal cell proliferation and analgesia. Recently, this receptor was found to be present in T- and antigen-presenting cells, where it regulates the proliferation and functions of these cells. The Y1 receptor requires the whole NPY(1–36) molecule for activation; therefore, peptidases that cleave N-terminal amino acids inactivate Y1 receptor-mediated activities (i.e., act as endogenous Y1 antagonists). The major enzyme in this group is dipeptidyl peptidase IV (DPPIV), which is present ubiquitously in endothelial, immune cells and macroglia, and forms NPY(3–36). The other known protease is aminopeptidase P (APP), expressed on vascular smooth muscle and other cells, which forms NPY(2–26).

On cleavage, C-terminal fragments of NPY can be detected in the circulation and remain active, but change their receptor affinities and become poor activators for or are inactive at Y1 while having a high affinity for the other, non-Y1 receptors. Up to five more receptors have been cloned so far: Y2, Y4, Y5, and Y6. All of them are Gi-coupled. The remaining Y3 receptor has been characterized pharmacologically as the one that is NPY-preferring, that is, not activated by PYY; it has not been cloned. Although a number of tissues appears to display such a pharmacological profile, for example, the brain stem, this receptor is now considered nonentity.

Y2 receptors are the predominant neuronal presynaptic receptors in the brain and the periphery, where they inhibit neurotransmitter release. They act as NPY's own autoreceptors as well as inhibiting the release of NE and other neurotransmitters such as ATP. Interestingly, the Y2-mediated inhibition appears to increase at a higher frequency of neuronal activity. In the central nervous system, the inhibitory effects of Y2 receptors on NPY release often result in the inhibition of either Y1- or Y5-receptor-mediated responses. For that reason, Y2 receptor antagonists are now emerging as the preferred drugs for several mental disorders (e.g., anxiety and depression) because they allow targeted therapy to the sites of overstimulation and limit other widespread effects in the brain that Y1- or Y5-receptor antagonists could otherwise have.

Postsynaptic Y2 receptors have also been found in the blood vessels, where they regulate the vascular tone; these changes range from vasoconstriction to vasodilation and angiogenesis, depending on the release from the endothelium of secondary mediators, such as nitric oxide. Whereas Y1 and Y2 receptors are constitutively expressed in unstimulated tissues, Y5 receptors appear not to be expressed alone and are often induced on demand in response to cellular stress. Initially, the Y5 receptor has been cloned from the hypothalamus as a feeding receptor, mediating NPY's action to stimulate appetite. It shares with the Y1 receptor the same promoter sequence, although read in the opposite direction, and appears to be regulated in concert with the Y1. Recently, we reported that Y5 is inducible in endothelial and vascular smooth muscle cells by tissue ischemia and injury.

Sources of NPY Release in the Circulation

The role of NPY as a mediator of nonadrenergic sympathetic transmission has been well documented by several groups in early studies both *in vitro* and *in vivo*. It became generally accepted that the peptide is released by more intense or prolonged nerve stimulation than that required for the release of NE. The

differential release of NPY and NE has also been seen using various stressors as stimulators of sympathetic nerve activity. A plausible explanation has been the different distribution of these neurotransmitters in the secretory granules, with NPY being stored in the large and NE in the small dense core vesicles. Although NE is also stored in the large vesicles, those constitute only a minor portion of the overall adrenergic pool and hence are minor determinants of the amount of released NE. However, recent studies with cultured cells have demonstrated that NPY is secreted from neuronal and non-neuronal cells at an even lower intensity of stimulation than seen before. In addition, NPY has been found to exert biological effects at extremely low concentrations (submolar to picomolar range). The *in vivo* detection of such minimal amounts of the peptide is clearly beyond the limits of currently used methods and, hence, cannot be documented at the present time.

Radioimmunoassays and enzyme-linked immunoassays have been the most commonly used methods for detecting circulating and tissue-derived NPY. Plasma NPY-immunoreactivity (NPY-ir) was found to vary among the species, with human and porcine levels being the lowest (picomolar range). Rats and some mice havie higher NPY levels (submolar to nanomolar range) due to the expression of NPY extraneuronally in megakaryocytes and platelets. In these animals, NPY-containing platelets become a transport system delivering the peptide to the sites of platelet activation. Although healthy humans do not have NPY in platelets, some diseases such as atherosclerosis appear to induce its expression. The significance of these findings remains to be determined; however, animal studies suggest that platelets, as well as NPY, play a critical role in vascular restenosis after balloon angioplasty. Interestingly, the levels of NPY-ir in the systemic arterial and the coronary sinus plasma increase during angina due to either myocardial ischemia or panic attack, indicating that intracardiac rise in NPY may contribute to vascular occlusion and be involved in thrombus formation, vasosopasm, and vascular remodeling.

The third source of circulating NPY is the endothelium; however, a controversy still exists about whether these cells actually secrete the peptide into the bloodstream. NPY is an autocrine growth factor and stimulates endothelial cell migration, proliferation, and differentiation into capillaries by an action mediated primarily by Y2 and Y5 receptors, with an active role played by DPPIV (components of NPY's angiogenic system are expressed in these cells). Ischemia and cellular stress both upregulate NPY, Y2 and Y5 receptors, and DPPIV.

Stress-Induced Release of NPY from Peripheral Sympathetic Nerves

The highest plasma NPY-ir levels in humans were noted in newborns after vaginal delivery. This is probably due to the stress of delivery; babies born by cesarean section have significantly lower peptide levels. With aging, circulating NPY levels in humans and in animals decrease, but stress continues to stimulate its release. Various stressors have been studied and found to exert differential effects on plasma NPY, by itself and in relation to catecholamines. Of those, the strongest stimulus for NPY release appears to be exposure to cold temperature and strenuous exercise, particularly when combined with hypoxia. Interestingly, there is a gender difference in the response to cold and/or treadmill exercise in both humans and rodents. Although there is no sexual dimorphism in the basal circulating NPY levels, males increase their plasma NPY-ir levels more than females following both stressors. In rats and in humans, this does not appear to be due to estrogens because ovariectomy and estrogen replacement has no effect on cold- and exercise-induced NPY responses. Subsequent studies in gonadectomized rats confirmed that the gender difference in the NPY system is due to testosterone, which strongly upregulates NPY gene expression. Sex differences in stress-induced plasma NPY responses were paralleled by greater hypertensive effects of stress in males, implying that the peptide may play a role in essential hypertension, which is more frequent in men than in premenopausal women.

Interestingly, other commonly used laboratory stressors such as immobilization stress in rats or mental arithmetic in humans failed to yield measurable changes in circulating NPY-ir levels using current analytical methods. Interestingly, a rise in plasma NPY levels in humans during bicycle exercise is augmented by the blockade of α-2-adrenergic receptors, indicating that NPY release is under physiological (presynaptic) inhibition by these receptors. The depletion of catecholamines from the sympathetic storage vesicles by reserpine was also shown to increase the NPY-dependent component of the neurogenic response to nerve stimulation (after a transient fall in blood pressure due to the removal of adrenergic vascular tone). Ganglionic blockade with chlorisondamine eliminated the NPY release and resulted in a complete loss of sympathetic tone. This experimental protocol elegantly demonstrated that although NPY may not be the primary neurotransmitter maintaining normal vascular tone and blood pressure, it is the alternative pathway to adrenergic transmission and is one that can either compensate for the loss of it or

amplify or dominate over the adrenergic response. An example of the former may be sympathetic dysautonomia, in which patients have impaired adrenergic transmission but are able to compensate for it; an example of the latter may be situations in which NPY signaling is selectively enhanced by cold exposure and a high-fat and high-sugar diet.

Chronic Stress, NPY, and Vascular Remodeling after Balloon Angioplasty

Although the examples so far illustrate the ability of the body to secrete NPY in response to acute stimulation of the sympathetic nerves, the role of this peptide neurotransmitter as a stress mediator can be understood best in chronic stress. In several diseases in which sympathetic activity is increased, circulating NPY-ir levels are markedly elevated. Some of the highest levels in humans were found in congestive heart and renal failure, in which they strongly correlated with and might be predictors of end-organ damage.

The actions of NPY also differ when it is applied chronically. In the cardiovascular system, acutely, NPY causes a slow onset, long-lasting vasoconstriction and increase in blood pressure at nanomolar concentrations consistent with the levels yielded by acute intense stress, for example, the immersion of a hand in ice water for humans or standing in ice-cold water for rodents. *In vitro*, such NPY concentrations applied for 1–14 days to cultured vascular smooth muscle and endothelial cells stimulates their migration, proliferation, differentiation, and release of secondary growth factors, leading to sustained cell growth. In these *in vitro* conditions, NPY also acts at lower than vasoconstrictive concentrations (submolar to picomolar range) and acts as an autocrine and paracrine growth factor. In *ex vivo* cultured aortic rings, NPY is potently angiogenic and leads to the formation of long vascular sprouts, resembling normal vessels. For some cells, such as neuroblastomas (tumor cells derived from sympathetic neurons), NPY is essential as a cell growth factor, whereas for others, such as Ewing's sarcoma, NPY is apoptotic. Thus, growth-modulatory actions are mediated by its multiple receptors and the final effect depends on the type of receptors and cell-specific signaling pathways. Whereas the growth-promoting effects of the peptide on vascular smooth muscle are mediated by Y1 receptors, NPY's angiogenic activities are mediated primarily by Y2 receptors and require the conversion of NPY(1–36) to NPY(3–36). What is the significance of these actions *in vivo*? Would the chronic stress-induced release of NPY mimic the peptide's effects seen *in vitro* and, if so, which ones?

We approached these questions using cold-water stress (animals standing in 0.5–1 cm of ice-cold water at the bottom of their cages) combined with experimental conditions known to trigger vascular remodeling. In rats, balloon angioplasty-induced arterial injury or intermittent chronic cold by itself within 2 weeks increased the platelet NPY-ir levels, but only cold stress elevated the plasma NPY. This suggested that, although plasma NPY responses are probably due to sympathetic nerve activation, platelet NPY may reflect changes originating in the bone marrow, in the megakaryocyte synthesis of NPY. This notion has been supported by studies using mice expressing or not expressing NPY in megakaryocytes. Both rats and mice, which have additional source of NPY in the bone marrow, developed neointima in the angioplasty-injured vessels, whereas mice deficient in megakaryocyte/platelet NPY were resistant to vascular injury and did not develop neointima. Whether platelet NPY is critical for neointima formation in rodents, and in humans, is currently being investigated.

NPY that is exogenous or released from sympathetic nerves by cold stress, on the other hand, augments neointima in all rodent strains, with or without platelet NPY, although the extent of the amplification is significantly greater in the former. In angioplasty-injured rat carotid arteries, NPY and stress alike induced the occlusion of the angioplasty-injured vessel with a lesion, which contained neovascularized neointima, thrombus, matrix, lipids, and increased invasion of macrophages – the elements of advanced atherosclerotic lesion. This occurred in an animal devoid of any lipid abnormalities within only 2 weeks and was completely prevented by a selective Y1 receptor antagonist. How it is possible for an antagonist of only one receptor to block such a complex process as restenosis and atheroma-like lesion formation is indeed puzzling. The critical role of Y1 receptors, however, is better understood when we consider the environment into which NPY is released following sympathetic nerve activation and vascular injury.

Critical Role of Y1 Receptors in Immune System and Angioplasty-Induced Restenosis

Chronic stress exerts direct effects not only in the cardiovascular system but also on the immune cells, commonly manifested as a greater incidence of infection during periods of stress. Although many mediators, including catecholamines and glucocorticoids, contribute to this effect, the role of NPY has not been considered until recently. Wheway et al. recently

reported that NPY and its Y1 receptors are essential for the normal functioning of T cells. The direct stimulation of Y1 receptors on T cells leads to the suppression of their Th1-type response and decreased secretion of interferon-1γ and tumor necrosis factor-α. For that reason, Y1 receptor agonists are protective against Th1-mediated inflammatory conditions such as experimental autoimmune encephalitis, a model of multiple sclerosis. However, NPY can also exert stimulatory actions on immune cells via the same type of Y1 receptors on antigen-presenting cells. The loss of that response early during development in germline Y1 receptor knockout mice led to the impaired functioning of antigen-presenting cells, reduced number of functional T cells, and, hence, lowered Th1-mediated responses. Paradoxically, Y1 receptor deficiency protects against colitis, which depends on Y1-mediated activation of antigen-presenting cells.

The central role that the Y1 receptor plays in the immune system is due to a widespread distribution of that receptor, not only in vascular smooth muscle cells but also in all immune cells, including T cells, macrophages, neutrophils, and eosinophils. NPY activates monocytes and stimulates their phagocytosis, migration, and proliferation. This is likely to have implications outside infectious or autoimmune diseases because inflammatory responses are now considered important pathogenic mechanisms in many other conditions, including atherosclerosis, ischemic angiogenesis, and obesity. Thus, we can imagine that the stress-induced activation of the sympathetic nerves during vascular injury may not only exert direct effects on vascular smooth muscle and platelets but may also activate immune cells residing in the vascular wall and stimulate their transmigration from the bloodstream. Because all these cells express Y1 receptors, it is likely that nerve-released NPY starts a cascade of events by activating this receptor on cells adjacent to the nerve terminals. In addition, the Y1 receptors are upregulated by angioplasty, and this increase is further amplified by stress; as a result, Y1 receptors become widely expressed in neointima, macrophages, and other immune cells.

Because the Y1 receptor antagonist prevented the entire process of restenosis in rodent models of balloon angioplasty, we can also predict that Y1 deficiency is protective against stress-induced restenosis and, that Y1 receptor antagonists may be a new class of drugs for this condition. Our studies also suggest that stress is a risk factor for restenosis, facilitating the formation of atheroma and vascular hyperplasia even in the absence of hyperlipidemia. Such striking changes as those induced by stress have never been previously induced in animals without genetic manipulations of lipid metabolism. Although more studies are needed to establish the clinical relevance of these findings, two lines of evidence strongly suggest that high NPY levels, and NPY release by stress, play a role in human atherosclerosis. First, an NPY signal peptide polymorphism, Leu7Pro, a common gain-of-function mutation occurring in northern European populations was found to lead to higher plasma NPY levels in response to stress, along with hyperlipidemia and accelerated atherosclerosis. Second, the induction of NPY in platelets was recently discovered in patients with atherosclerosis, similar to what previously was reported to occur in patients with severe depression. Current studies in our lab are being conducted to confirm whether stress in humans accelerates atherosclerosis by releasing NPY from the sympathetic nerves and/or platelets.

Chronic Stress and Angiogenesis: Implications for Ischemic Revascularization and Adipogenesis

If stress releases NPY from nerves and/or platelets, what determines whether the peptide exerts vasoconstrictive vasoproliferative effects leading to vascular hyperplasia via its Y1 receptors or is angiogenic leading to the increased vascularization of the tissue via its Y2 receptors (Figure 1)? We addressed this question first using a model of ischemic limb, in which

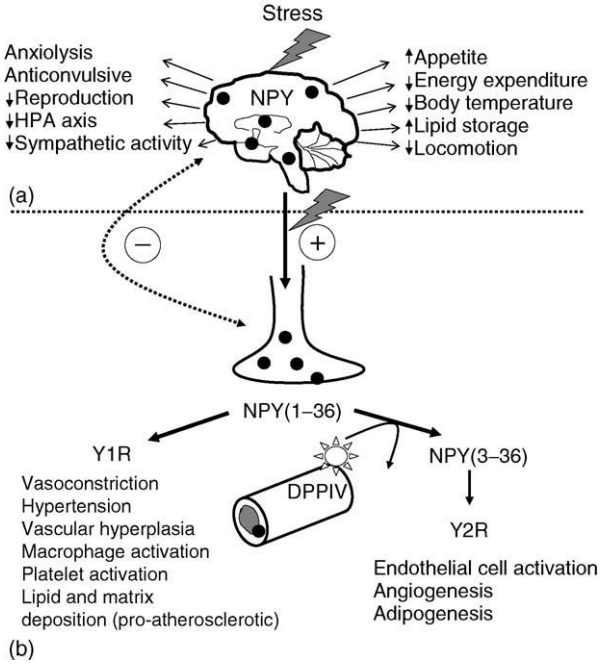


Figure 1 Effects of stress on central and peripheral NPY systems. a, Central nervous system effects; b, peripheral effects in sympathetic nerves. DPPIV, dipeptidyl peptidase IV; HPA, hypothalamic-pituitary-adrenal; Y1R, NPY receptor 1; Y2R, NPY receptor 2.

spontaneous revascularization is incomplete and produces a deficit in blood flow. Limb ischemia induced by unilateral femoral artery occlusion, like angioplasty described previously, elicited a local release of NPY from the sympathetic nerves. However, in contrast to the effects of angioplasty on the vascular wall, tissue ischemia also stimulated NPY synthesis in the endothelium and shifted the expression of its receptors in the ischemic limb toward Y2 and DPPIV, the elements of NPY's angiogenic system. The local administration of NPY at the site of ischemic skeletal muscles amplified these responses and, in addition, also induced Y5 receptors while augmenting angiogenesis and arteriogenesis in the limb. Exposure to cold, intermittently for 2 weeks, however, did not mimic NPY's actions and actually inhibited ischemia-mediated changes in NPY receptor expression. Although these studies are still in progress, the data suggest that NPY's angiogenic receptor system, NPY-Y2/Y5-DPPIV, is not sufficiently activated during stress. Yet unidentified factors appear to inhibit the NPY's angiogenic switch and prevent the peptide from fully expressing its angiogenic capabilities in the ischemic tissues. Given the fact that Y2 and Y5 receptors and DPPIV are expressed in the ischemic skeletal muscles, the administration of Y2/Y5 receptor agonists may become useful as a revascularization therapy. Alternatively, the activation of DPPIV, the converting enzyme, could serve as a way to shift stress-induced NPY release away from vasoconstriction and vascular hyperplasia via its Y1 receptors and toward Y2/Y5-mediated angiogenesis.

Such a switch appears to occur spontaneously in the white adipose tissue under the conditions of a high-fat and high-sugar diet; under this condition, stress causes a marked upregulation of NPY, Y2, and DPPIV. This leads to the activation of angiogenesis, which then promotes further tissue growth and, thus, may contribute to the development of adiposity and obesity. Previously, others have proposed using anti-angiogenic therapy for the treatment of obesity, but the physiological mechanisms specific to fat have not been uncovered. This new area of research should shed light on the pathogenesis of obesity, which has become an epidemic concurrently with the rise in the levels of daily stress. The relation between the two processes, stress and obesity, although anecdotally acknowledged, has never been established scientifically.

Summary and Conclusion

In addition to the roles of NPY in the cardiovascular system described here, the peptide exerts pleiotropic actions on essentially every tissue and the regulatory system of the body. In the brain, NPY is potently anxiolytic, sedative, and anticonvulsive and inhibits locomotor activity – actions consistent with the anti-stress effects rather than the pro-stress actions that it exerts in the periphery (Figure 1). Those central actions have implicated the peptide in anxiety and mood disorders, in which NPY levels in the cerebrospinal fluid are decreased. The peptide is also one of the most powerful stimulants of food intake, particularly of carbohydrates, resembling stress-induced cravings for sweets, and thus plays a role in appetite disorders, both obesity and anorexia. At the same time, NPY modulates a number of hormonal systems, from the hypothalamic-pituitary-adrenocortical and hypothalamic-pituitary-gonadotropin axes to the secretion of growth hormone and insulin, slowing down puberty and suppressing reproductive functions (Figure 1). In addition to its peripheral actions, centrally, NPY also inhibits lipid metabolism and energy expenditure, lowers body temperature, and promotes fat storage (Figure 1).

The two sites of NPY's actions, the brain and the periphery, have been studied so far in separation and little is known about their reciprocal relationship. It appears, however, that, as with many other neurohormonal systems, central and peripheral NPY are also regulated by negative feedback loops, with central NPY being often inhibitory for peripheral NPY. A primary example of such a reciprocal inhibitory relationship is the inhibitory effect of hypothalamic NPY on peripheral sympathetic activity, lowering plasma NPY levels. In conditions such as posttraumatic stress disorder, resulting from severe trauma, the balance between the central and peripheral NPY responses appears to be disregulated at both brain and the peripheral levels. The regulation of the cardiovascular system by NPY in this and other stressful situations is clearly only a part of the patterned body responses to stress and the patterns that vary with different stressors.

The integrated response of the NPY system in many ways resembles the body's adaptation to hibernation. Several lines of evidence support that hypothesis. First, cold temperature is one of the most potent stimuli for NPY release. Second, in preparation for hibernation, an animal requires cardiovascular adaptation, increased blood pressure and heart rate, to allow it to search for food; such a response is provided by peripherally released NPY. Third, in the presence of carbohydrate-rich food, the NPY stress response also promotes lipid storage, which an animal preparing for hibernation needs. Once this is accomplished, sympathetic activity is expected to fall and the animal's bodily functions to slow down. NPY's central activities support such actions as well,

with its effects of decreased body temperature, energy expenditure, and reproductive functions.

The final intriguing support of the hibernation hypothesis is suggested by genetic epidemiology. A Leu7-to-Pro7 polymorphism in the NPY signal peptide has been found to be common in northern European populations. Interestingly, there is a strong north-to-south gradient of this mutation, which is absent in Africa, Asia, and South America (unless in descendants of northern Europeans). The Leu7Pro polymorphism associates with increased NPY responses to stress, hyperlipidemia, and obesity – which at some time could have afforded an evolutionary protection, improving survival in cold temperatures. Unfortunately, there are also unwanted effects of this mutation. It is linked to accelerated atherosclerosis, diabetic retinopathy and nephropathy, and alcoholism – perhaps the price that humans paid for not adjusting to the rules of nature, eating and drinking too much while exercising too little all year long.

The importance of NPY in physiology was questioned by the studies of the Palmiter group who showed that life without NPY, in NPY-knockout mice, is apparently normal. The mice eat, develop, and breed normally and do not have any specific phenotype. They do, however, exhibit a lower threshold for seizures and have higher preference for alcohol – possibly offering a new model for epilepsy and alcoholism. Yet the true contribution of NPY to animal physiology can only be appreciated when the adaptation to chronic stress is considered. When laboratory animals are housed in a constant environment, without any stress, their life may seem normal. However, if NPY-knockout mice are subjected to stress, particularly chronically and in combination with poor diet, their ability to develop proper cardiovascular, immune, and metabolic adaptations might be impaired. In years to come, more studies will be required to understand the full implications of NPY's release and activities in all of the body's sources: the brain, peripheral sympathetic nerves, and the enteric system, as well as its extraneuronal expression in platelets, endothelium, and immune cells.

Further Reading

Chronwall, B. and Zukowska, Z. (2004). Neuropeptide Y, ubiquitous and elusive. *Peptides* **25**, 359–363.

Dumont, Y. and Quirion, R. (2006). An overview of neuropeptide TY: pharmacology to molecular biology and receptor localization. In: Zukowska, Z. & Feuerstein, G. Z. (eds.) *NPY family of peptides in neurobiology, cardiovascular and metabolic disorders: from genes to therapeutics*, pp. 7–35. Basel: Birkhauser Verlag.

Dutton, M. A., Lee, E. W. and Zukowska, Z. (2006). NPY and extreme stress: lessons learned from posttraumatic stress disorder. In: Zukowska, Z. & Feuerstein, G. Z. (eds.) *NPY family of peptides in neurobiology, cardiovascular and metabolic disorders: from genes to therapeutics*, pp. 213–247. Basel: Birkhauser Verlag.

Erickson, J. C., Clegg, K. E. and Palmiter, R. D. (1996). Sensitivity to leptin and susceptibility to seizures of mice lacking neuropeptide Y. *Nature* **381**, 415–418.

Heilig, M. (2004). The NPY system in stress, anxiety and depression. *Neuropeptides* **38**, 213–224.

Li, L., Rylander-Johnsson, A. C., Abe, K., et al. (2005). Chronic stress induces rapid occlusion of angioplasty-injured rat carotid artery by activating neuropeptide Y and its Y1 receptors. *Arteriosclerosis Thrombosis and Vascular Biology* **25**, 2075–2080.

MacNeil, D. J. and Kanatani, A. (2006). NPY and energy homeostasis: an opportunity for novel anti-obesity therapies. In: Zukowska, Z. & Feuerstein, G. Z. (eds.) *NPY family of peptides in neurobiology, cardiovascular and metabolic disorders: from genes to therapeutics*, pp. 143–157. Basel: Birkhauser Verlag.

Malmstrom, R. E. and Lundberg, J. M. (1997). Neuropeptide Y in sympathetic nerves: evidence for Y1 receptor mediated vascular control. In: Grundemar, L. & Bloom, S. R. (eds.) *Neuropeptide Y and drug development*, pp. 41–49. London: Academic Press.

McAuley, M. A., Chen, X. and Westfall, T. C. (1993). Central cardiovascular actions of neuropeptide Y. In: Comers, W. F. & Wahlestedt, C. (eds.) *The biology of neuropeptide Y and related peptides*, pp. 389–418. Totowa, NJ: Humana Press.

Pesonen, U. (2006). Genetics of the NPY family of peptides and their receptors. In: Zukowska, Z. & Feuerstein, G. Z. (eds.) *NPY family of peptides in neurobiology, cardiovascular and metabolic disorders: from genes to therapeutics*, pp. 247–271. Basel: Birkhauser Verlag.

Wheway, J., Mackay, C. R., Newton, R. A., et al. (2005). A fundamental bimodal role of neuropeptide Y1 receptor in the immune system. *Journal of Experimental Medicine* **202**, 1527–1538.

Zukowska, Z. (2005). Atherosclerosis and angiogenesis: what do nerves have to do with it? *Pharmacological Reports* **57**, 229–234.

Zukowska, Z., Grant, D. S. and Lee, E. W. (2003). Neuropeptide Y: a new angiogenic mechanism for ischemic angiogenesis. *Trends in Cardiovascular Medicine* **13**, 86–92.

Zukowska-Grojec, Z., Shen, G. H., Capraro, P. A., et al. (1991). Cardiovascular, neuropeptide Y, and adrenergic responses to stress are sexually differentiated. *Physiology and Behavior* **49**, 771–777.

Zukowska-Grojec, Z. and Wahlestedt, C. (1993). Origin and actions of neuropeptide Y in cardiovascular system. In: Colmers, W. F. & Wahlestedt, C. (eds.) *The biology of neuropeptide Y and related peptides*, pp. 315–388. Totowa, NJ: Humana Press.

Stress-Induced Neurotoxicity *See:* Glucocorticoids – Adverse Effects on the Nervous System.

Stroke, Stress and *See:* Brain Trauma.

Successful Adaptation to Extreme Stress *See:* Multiple Trauma; Sepsis, Acute Respiratory Distress Syndrome and Glucocorticoid Resistance; Acute Stress Disorder and Posttraumatic Stress Disorder; Acute Stress Response: Experimental.

Suicide Terrorism, Genesis of

A Speckhard
Georgetown University Medical Center, Washington, DC, USA

© 2007 Elsevier Inc. All rights reserved.

The History of Suicide Bombing
The Success of Suicide Terrorism
The Lethal Cocktail of Suicide Terrorism
Sponsoring-Groups Motivations
Ideologies of Suicide Terrorism – Religion and Cosmic Warfare
Individual Motivations
Adopting a Terrorist Ideology as Psychological First Aid
Societal Support for Suicide Terrorism

Glossary

Dissociation According to the American Psychiatric Association, the essential feature of a dissociative disorder is "a disruption in the usually integrated functions of consciousness, memory, identity or perception of the environment" – American Psychiatric Association (1994) *Diagnostic and statistical manual of mental disorders (4th ed.)*. Washington, D.C. In the case of suicide terrorism, this means an emotional barrier is unconsciously erected walling off the negative emotions generated by choosing to die in this manner and may even include compartmentalizing the event from one's ambitions and daily life. One bomber for example, when planning his attack, suggested that he could not carry a bomb until after his university exams – suggesting that while on the one hand he acknowledged that he was going to his death he was able separate this reality so completely from his mind that he still felt the need to complete his exams prior to going to explode himself.

Jihad Is an Islamic word and makes reference to the teachings of the Prophet Muhammed as they are recorded in the Koran and Sunnah, in which believers are taught to struggle on two levels: the first and most important being referred to as the greater jihad in which an individual strives inwardly in an effort to resist evil and live a pure life according to the tenets of Islam. The lesser jihad involves a political or military struggle to further and/or defend Islam and Islamic lands and people. In the current global context certain ideologies promoting terrorism are referred to by some as promoting militant jihad. For example the US Department of Justice has referred to jihad as the use of violence, including paramilitary action against persons and governments deemed to be enemies of a fundamentalist version of Islam.

Posttraumatic stress disorder (PTSD) Is a person's individualistic response to a traumatic event (see psychological trauma below) in which the following responses occur: an inability to integrate and respond well to a traumatic event on a cognitive, emotional and/or on a sensory level resulting in 1) reexperiencing behaviors including flashbacks, nightmares and unbidden and upsetting thoughts of the event; 2) avoidance of reminders of the event including feelings of social alienation; 3) bodily arousal in response to reminders of the event including shakiness, sweaty palms, loss of concentration, irritability, trouble falling and/or staying asleep; 4) loss of function in life domains including work, intimate and family relationships etc. and 5) all of these symptoms lasting for more than thirty days following the event. Shorter, (less than thirty days following the trauma) but equally disabling and intense, dissociative responses to psychological traumas which at times might transition into PTSD, are labeled as acute stress disorders.

Suicide or human bomber A suicide bomber as anyone who goes so far as to strap on a bomb, drive a vehicle filled with explosives to a target, or who otherwise attempts to detonate an explosive device on an airplane, in a subway, train, car, or elsewhere with the aim of dying to kill – irrespective of whether or not the bomber actually died in the attack or was successful in detonating – as that is often not within the bomber's control. We take the fact of being to the point of willingly strapping on a bomb or other type of improvised explosive device or driving a vehicle loaded with explosives to a target as sufficient evidence of seriousness of the intent to suicide and see the end result, which is often out of the hands of the bomber as less meaningful than the intent implied by these actions. An actor who arrives at a scene with a weapon, most often a gun, in order to shoot as many victims as possible prior to being killed himself may also be seen as a suicide terrorist, but is obviously not a human bomber. Psychologically the preparation and action of going on such a killing spree in which one's death is nearly guaranteed is very similar to that of a human bomber.

Terrorism Has many definitions but is generally viewed as political violence carried out by politically organized non-state actors/groups with the intent to destroy property, injure, take hostage and/or kill civilians and thereby intimidate and influence a wider witnessing audience and to effect political, religious or ideological change in both the targeted population and its government's responses; as well as to drum up support among the group's own sympathesizers. Terrorism relies heavily upon the mass media to amplify the psychological effect of attacks made on relatively small numbers of victims by relaying horrifying images to a much larger witnessing audience. There is considerable controversy over who is defined as terrorists, particularly in situations where non-state actors (i.e. freedom fighters, rebels or guerilla fighters) could be also seen as leading legitimate rebellions in occupied territories but who nevertheless use similar tactics as terrorists such as using civilians as shields and indiscriminately targeting civilian as well as military targets.

Psychological trauma Involves an experience in which an individual experiences a serious threat (real or perceived) to his own life or bodily integrity or witnesses death or threat of life or dismemberment in another and responds with overwhelming feelings of fear, horror and helplessness. In this experience the person is typically psychologically overwhelmed and unable to integrate his emotional, cognitive and sensory experiences in a normal manner and a peritraumatic dissociation may result immediately or in some time following the event. Acute or posttraumatic stress disorder may result.

Suicide terrorism is fast becoming one of the most pestilent global afflictions of the twenty-first century. As a terror tactic, it is one of the most lethal. Strategically it functions as a relatively cheap and effective means of upsetting the political, economic, and military situation of a region and has become one of the major threats to peacekeeping and peace-making efforts. Yet, as prevalence rates and death tolls from suicidal attacks increase, policy makers are still working in the dark trying to find the most effective policy responses to the emergence of this new and poorly understood security threat. Currently there is an extremely small empirical research database on which policy makers may base their understanding of the genesis of suicidal terrorism in order to work toward its prevention and eradication. Comprehending this growing threat and learning to combat it effectively on both the local and international levels is extremely relevant to current public policies aimed at promoting

peace and stability. This article briefly discusses the background of modern-day suicide terrorism, its migration around the world, and the tendency of modern-day terror groups to embrace it as a tactic. Drawing from field research in four distinct world regions, I give an in-depth analysis of the social and psychological factors in the genesis of human bombing.

Focusing on social and psychological factors within the sponsoring groups, the individual bombers, and the societies in which they exist, I have identified two main differing motivational sets on the level of the individual actors for the genesis of suicide terrorism. The first of these is trauma-based, occurs within zones of active conflict, and is often nationalistic, is viewed in terms of self and community defense, is expressive regarding attempts to mete out justice to the perceived enemy occupier, and includes acts of revenge by actors who are often so traumatized that they have become deeply dissociative and even refer to themselves as "already dead." The second motivational set applies to actors outside the active zones of conflict, who are nevertheless influenced by them (through the Internet, video footage, pictures, and propaganda) and who frequently develop a deep sense of secondary traumatization. The actors in the second motivational set are generally vulnerable to terrorist ideologies due to a sense of alienation, marginalization, lack of life meaning, and lack of positive identity and are usually recruited through exposure via the Internet or via close-knit family and friendship networks. In both motivational sets, it is important to recognize that individuals are generally not motivated to take part in suicide terrorism without a guiding ideology and that they are also generally in need of an organization to equip and guide them to carry out their acts (although a rare few have acted on their own; Argo, 2004; Speckhard, April 2005). They generally are more willing to take part in suicide terrorism when there is broad-based support for suicide terrorism in the sector of the community in which they find their sense of belonging.

The History of Suicide Bombing

Dying to kill is not new.[1] What is new is that, in modern times since the mid 1980s, the tactic of suicide terrorism has been repackaged and reborn. It has suddenly taken off like a wildfire spreading from Lebanon to Palestine, Sri Lanka, Chechnya, Morocco, Indonesia, Turkey, Afghanistan, and Iraq (where there is, on average, one suicide bomber per day); and now, sadly, it has spread even to those originating in Europe (i.e., the infamous shoe bomber, Richard Reid; the Mike's Place bombers in Israel; and Europeans going as bombers to Iraq and Afghanistan) and European-bred bombers now even targeting Europe itself. What combination of factors and conditions have made this tactic suddenly so popular that citizens of countries that have no history of suicide terrorism are suddenly willing to recruit themselves as human bombs, making this a main tactic of choice for many terror groups worldwide?

The Success of Suicide Terrorism

Looking back in recent history, we see that the current spate of modern-day suicide terrorism began in Lebanon in the 1980s when terrorists used suicide truck bombs to attack first the U.S. Embassy and later the barracks of the U.S. Marines and French troops, actions that led to the troops deciding to remove themselves from Beirut and position themselves offshore – this was viewed as a huge strategic victory by the terrorists.

Indeed the current epidemic of suicide terrorism is directly tied to this perception of its success by terror-sponsoring groups. As its use migrated around the world even many counterterrorism experts have credited suicide terrorism with derailing the Oslo peace accords, disrupting the peacekeeping and rebuilding efforts in Afghanistan and Iraq, impacting the election results in Madrid,[2] and drawing world attention and concern to political issues that may otherwise have been overlooked. Whether it is effective in achieving any real political gains outside the community in which it originates, that is, whether it creates

[1] I am indebted to Mia Bloom for coining this phrase, which she also uses as the title for her book Dying to Kill: The Allure of Suicide Terror (Columbia University Press, 2005). Suicide bombing as a strategic tactic is not a new phenomenon. In recent history, it appeared in the 1950s in Vietnam in the form of bicycle bombers who exploded themselves in cafes, killing the enemy occupiers who frequented them (Francois Gere, French counterterrorism expert, personal communication 2005). Looking back into the ancient past, many cite the scriptural account of Sampson as one of the first suicide terrorists; when blinded and chained to pillars, he decided to use his great strength to pull the columns down, collapsing the building on himself and those surrounding him. Similarly, some experts recall the Muslim Assassins and Jewish Sicari as the first suicide terrorists; both groups carried out assassination missions that were nearly always suicidal for the person carrying it out. Also Pakistani warriors in ancient times were known to run underneath the elephants carrying advancing troops to slit the bellies of the warring beast, bringing them down on themselves and killing both their enemy and themselves in the process.

[2] Some argue that the Madrid bombings did not involve suicide terrorists because the attackers left their bombs and detonated them from afar; yet threatened with imminent arrest, the bombers did explode themselves.

any real power base for those who employ suicide terrorism, is still, however, debatable.

Terrorist organizations thrive because of failed political solutions between conflicting parties, and their continued existence relies on their ability to change perceptions and impact the political process in favor of their constituent groups. When a much weaker group is pitted against a larger, more powerful, better financed and militarily equipped group, suicide terrorism becomes the tactic of choice because it is seen by the weaker group as strategically both efficient and more likely to lead to desired results than any other method of violence.

Suicide terrorism is a tactic that is:

- Inexpensive. It requires relatively simple, noncostly, and easily accessible equipment. The 9/11 attacks on the World Trade Towers and Pentagon cost approximately $400 000, the Bali bombings approximately $20 000 (NATO Defence College, 2004).
- Highly effective. The human bomber is essentially a smart bomb, going directly to his or her target and able to make adjustments even up to the last moments before detonation to avoid detection and to maximize the amount of damage done in the attack.
- Highly lethal. Suicide terrorism operations are the most lethal of nearly all currently used terror operations, and although they make up only a small fraction of the total of terrorist operations, they kill and wound the greatest number of victims.
- Extremely horrifying. The carnage caused by human bombers almost always makes the news; hence, by targeting only a small group of civilians the terror-sponsoring group can count on the media to amplify its effects, causing horror and dread to spread throughout a much wider witnessing audience.
- Nearly impossible to prevent. Once a bomber is equipped and on the way to the target, he or she is virtually impossible to stop (unless he or she can be talked out of the mission, which is rare) because he or she can always explode him- or herself on detection (Ganor 2005).
- Difficult to trace. The bomber, if successful, is killed in the attack and, unlike in other terror acts, needs no resources or risk dedicated to an escape plan; and once killed he or she cannot be caught and interrogated later, revealing who sent him or her.
- In endless supply, if the terror group's constituent population supports the use of this tactic. The main cost of suicide terrorism is the human beings who agree to sacrifice themselves on behalf of the terrorist cause. If this pool of individuals is large, the terror organization has a virtually endless supply of weaponry and can carry on a very strong war of attrition with a much more powerful enemy who may eventually make concessions to the terror-sponsoring group simply to end the campaign of bombings.

The Lethal Cocktail of Suicide Terrorism

Suicide terrorism originates from the confluence of four main aspects of suicide terrorism: the sponsoring group's motivations, the ideology that supports it, the individual motivations for enacting it, and the societal support for suicide terrorism. Without this confluence of factors, suicide terrorism ceases to exist (Argo, 2003; Hafez, 2004; Merari, 2003; Moghadam, 2003; Speckhard, 2005b; Speckhard and Ahkmedova, 2005). When all these factors are present, suicide terrorism becomes a tactic that can travel the world as a spark igniting and reigniting a massive fire that we can only hope to be able to extinguish before it consumes us in its flames.

Sponsoring-Groups Motivations

Organizations provide the means, methods, and group dynamic underlying suicide terrorism operations and often the ideology as well. However, terrorist ideologies should not be thought of as simple constructs of terror-sponsoring groups. They arise, in fact, from a complex mix of social and political circumstances, psychological and religious context, and interplay among the actions, sentiments, and rhetoric of terror groups, their constituency, and the perceived oppressor-enemy. When looking at the motivations on the organizational level, we must acknowledge that terror-sponsoring organizations are largely political in their motivations and resort to terrorism when other political solutions have failed and resort to suicide terrorism when a specific set of circumstances exist. Researchers of suicide terrorism know that it is nearly always used strategically by organizations and generally resorted to only when the enemy is much stronger and better equipped militarily. Likewise, it is often used in later stages of the conflict (Bloom, 2005); sometimes reflects an outbidding process for power among competing groups (Bloom, 2005); and finds a receptive base of support in areas where occupation occurs, particularly if the foreign occupier is of another religion than those occupied and is perceived as oppressive and unjust (Pape, 2005; Speckhard, 2005b; Speckhard and Ahkmedova, 2005; Speckhard and Akhmedova, 2006b). When

these circumstances exist, suicide terrorism may be deemed an effective choice by terror-sponsoring organizations for forcing concessions from its stronger enemy and thereby achieving its political goals.

Ideologies of Suicide Terrorism – Religion and Cosmic Warfare

Currently the strongest ideology supporting suicide terrorism makes use of distorted versions of Islam. As anyone who has studied the history of warfare knows, religion is often invoked to send warriors out to battle because believing that one is dying for a higher cause is highly motivating. Nearly all major faiths have in the past been used in this way and still are. Just as nearly all nations going to war today still make use of religious rhetoric when looking for popular support and to motivate their warriors, so too do many of today's terror-sponsoring groups (consider the rhetoric of the Crusades and the just war debates carried on in the West prior to the invasions of Afghanistan and Iraq). Although the loosely affiliated Al Qaeda/global Salafi terrorist groups do not represent legitimate nation-states, they claim that they are acting in behalf of a group of beleaguered people and that they are in a war, and their ideology is aimed to motivate warriors for the battle. The difference is not so much in the use of religion to garner popular support for acts of war and to motivate warriors but in the distortions of mainline religion to justify the tactic these warriors are motivated to adopt and the targeting of innocent civilians that their ideology justifies. We see in these current jihadist ideologies promulgated over the Internet and through underground networks the promotion of suicide terrorism – claiming it as the most effective way for these groups to triumph – the manipulation and misuse of treasured religious principles valuing martyrdom on behalf of Islam to motivate foot soldiers to self-recruit themselves for suicide missions against innocent civilian targets.

The suicide terror groups' uses of religion to motivate individuals to sign up to die makes strategic sense. Any believer of any faith who is persuaded of the following will act in extraordinary ways:

- He or she is in a cosmic battle (Juergensmeyer, 2000) involving apocalyptic forces
- that will eradicate either his side or the other, hence necessitating a war of defense.
- He or she should dehumanize and even demonize his or her enemies by seeing them as the enemies of God.
- In joining the group and taking on its values and teachings, he or she has learned the mind of God and is authorized to act in the battle by the will of God.
- The battle in which he or she is fighting is for sacred values (Atran, 2003).
- He or she may go to extraordinary means to eradicate and blot out those he or she sees as evil-doers – even innocent civilians – who he or she believes are standing in the way of and offending God's will.

It is important to state however that, even though groups that make use of distorted versions of Islam are currently highly effective in promoting this cosmic warfare ideology to endorse the tactic of suicide terrorism, this does not mean that Islam itself is the problem. If it were, we would not have seen the Liberation Tigers of Tamil Eelam (LTTE) and the Popular Front for the Liberation of Palestine (PFLP), which are mainly composed of Marxist atheists, making use of this same tactic. In those cases, they made no reference to Islam, or to religion at all, but relied on charismatic leadership, deep-seated anger over nationalistic concerns and injustices, and the hope of becoming heroes for their cause to motivate human-bombing recruits. Likewise, many cults have made similar use of non-Islamic religions to induce an apocalyptic vision of the world in their believers, who then become willing to endorse violence and even kill and die in order to bring it about; these include the Aum Shinryko, the Japanese Hindu-related cult responsible for the sarin-gas poisonings in the Tokyo metro; the People's Temple cult that followed Jim Jones, whose members made a suicide pact and killed themselves and their families, resulting in over 900 deaths; and the Army of God, the Christian-based abortion clinic bombers.

So, although Islam is currently being misused to promote suicide terrorism, it is not in itself the problem. The problem is the powerful links that terrorists groups are able to make among individual motivations to self-sacrifice, societal circumstances leading to despair and defiance, and a hijacked version of Islam that, playing on sacred scriptures, promotes human sacrifice on behalf of the group. Any mainline religion can be distorted and used to motivate for suicide terrorism by framing it as an ideology in support of self sacrifice in a cosmic war of defense for the true faith. Currently, the main ideology in use among the most active suicide terrorist groups (i.e., the loosely affiliated Al Qaeda/global Salafi and other nationalistic militant jihadist-linked groups) is a hijacked version of Islam that calls for would-be martyrs from around the world to sacrifice themselves on behalf of a worldwide or nationalist jihad.

Individual Motivations

But why would anyone answer the call? Why do most of the main terror groups in Palestine, Lebanon, Iraq, Afghanistan, Chechnya, and elsewhere state that they have an endless supply of recruits – such as the Chechen female bomber who spoke on a prepared video broadcast during the Dubrovka/Nord Ost theater takeover in Moscow, where 800 hostages were held for 3 days, stating, "Even if we are killed, thousands of our brothers and sisters will come after us ready to sacrifice themselves" (Jazeera, 2002). Certainly there must be something much more powerful behind this than simply an ideology to make an individual cross the line from normal daily life to strapping on a bomb or getting into an airplane, car, or truck intent on exploding him- or herself in the middle of a crowd of civilians. What are the factors that cause these changes to occur within a human soul?

Although the answer is multifaceted and impossible to answer in this short article, our research speaking to bombers, would-be bombers, senders, the family members and close acquaintances of bombers, and their hostages points to two main motivational sets (Speckhard, 2004, 2005b, 2005c, April 2005; Speckhard & Ahkmedova, 2005, 2006; Speckhard & Akhmedova, 2004, 2006a, 2006b; Speckhard, Tarabrina, Krasnov, & Akhmedova, 2004; Speckhard, Tarabrina, Krasnov, & Mufel, 2005a, 2005b): trauma-based motivations and alienation-based motivations.

Trauma-Based Motivations

The first motivational set is a trauma-based and occurs within zones of active conflict. It is often nationalistic, viewed in terms of self and community defense, and expressive regarding meting out justice to the perceived enemy-occupier, and it often includes acts of revenge. The individuals motivated within this set have witnessed firsthand and over the television their neighbors, family members, and loved ones being killed by what they view as an occupying force. Many have grown up witnessing countless acts of violence and, as a result, have not developed normally and often suffer from posttraumatic stress and dissociative disorders. Many have lost jobs, educational opportunities, been humiliated, and often struggle for basic daily needs and security. Although the majority of traumatized individuals in conflict zones will not become suicide bombers even if invited to do so, an extremely small group will become vulnerable to terrorist ideologies that promote this tactic. Feeling constantly agitated by traumatic flashbacks; unable to avoid daily reminders of their traumatic losses; and feeling in constant danger, bereaved, angry, and impotent, these individuals ultimately become so dissociative (i.e., separated from normal thoughts, perceptions, and emotions) and emotionally numb that they often refer to themselves as "already dead." Actually, for them dying is no longer a feared outcome; they have already psychologically and emotionally numbed themselves to human suffering, yet it keeps mercilessly and painfully intruding into their thoughts, so death may seem as a welcome release. Embracing it and exerting some control over when and how it occurs is sadly welcomed by these traumatized individuals.

In response to threat, a normal person reacts in a fight-or-flight response and will only move into dissocative defenses (i.e., shutting down the normal features of consciousness such as emotions, logic, and memory and sometimes physically freezing as well) when the threat becomes overwhelmingly horrific, terrifying, and life-threatening. When this distinction was explained to him, Al Aqsa Martyrs' Brigade sender of suicide bombers Zecharia Zubeidi explained, the individuals he equips to become human bombs are, according to his observations, caught inflexibly over long periods of time in a dissociative mode. He stated that, whereas combatants are flexible in their responses to violent conflict, these victims of the conflict are not. "They are completely different than us [fighters]. They have only one decision. We have many options. The thought of running away is always available. We can go and shoot." In contrast, he described the martyrs as locked into an inflexible dissociative mode caused by traumatic stress and the one decision that comes from it. He stated, "They get flashbacks all the time and for them death is a mercy. . . . For the martyr all the cells in his mind are dead except for one." According to Zubeidi, suicide bombers are in too much psychic pain to find another way to cope and become totally fixated on carrying out what they view as acts of community defense, expressions of pain, and enacting justice in response to "all that they have seen." Explaining their psychological inflexibility in reference to his own dissociative states, which occur from time to time (also in response to traumatic experience), he stated, "When I feel this way I stay there one or two hours, but that one [a bomber], after all that he has observed there is only that one thing [i.e., to end his life on behalf of the community]."

Suicide bombers are often as well or better educated and less poor than their peers (Atran, 2004; Merari, 2003), and in other circumstances they might have been leaders in their communities. They are acutely sensitive to their own suffering and

that of those around them and wish to make a difference, but, similar to normally depressed and suicidal people, they see only very limited avenues of action. They want to escape their psychic pain but to do so honorably and to use their lives – even if it means dying – to help their communities. They are uniquely vulnerable to an ideology that promises that they will be heroes for the cause and that they can make a difference in the sociopolitical situation faced by their communities. They believe that their deaths are only a doorway to a better place and that by dying in self-sacrifice they can change things both now and in the afterlife, reuniting with those who have gone before and later bringing with them relatives who they left behind. Vulnerable and in pain, they succumb to an ideology that seduces them into sacrificing themselves for what they believe is a greater cause. Just like us, they hope for a more just world in which human dignity and rights will be upheld; however, unlike us, they have been deluded into believing that their acts of killing innocent civilians might bring this into being and for this they sacrifice themselves.

Alienation, Marginalization, Loss of Identity, Secondary Traumatization, and Desire for Life Meaningfulness, Belonging, and Heroism

The second motivational set involves actors not living in conflict zones and is more complex. Whereas we might understand an individual who has seen his or her family member killed in front of his or her eyes and feels his or her country has been occupied seeking revenge, we must question what can possibly motivate Europeans, Turks, Moroccans, Uzbeks, and others join such groups ultimately agreeing to die in order to kill others? In their cases, the main motivational set appears to involve vulnerable actors who are exposed through kin and friendship groups (Sageman, 2004) or through the Internet and informal recruiting to other individuals within a terror network.[3] In their cases, these individuals are often marginalized, frustrated, and without hope in their societies. In Europe, they are often first-, second-, or third-generation immigrants or converts to Islam who feel deep sympathy and even kinship (i.e., as Muslim brothers) for those in conflict zones. Within the migrant community, there is often a deep sense of alienation and no secure sense of identity and belonging in either their country of origin and, more importantly, their host culture. Facing discrimination, often well educated but facing poor job prospects, lacking positive identity and a sense of life's meaning, and having little else to make of themselves, once they are exposed to terrorist ideologies they are attracted to the appeal to become heroes for a cause.

In these cases, the sponsoring organization nearly always makes use of five powerful motivators. The first two are the ideas of belonging and identity – belonging to an important cause and group – and of taking on a heroic identity. The third is the use of pictures and graphic video footage of conflict zones that are shown to the potential recruit and interpreted as atrocities against innocent victims – mainly featuring Chechen and Palestinian suffering, but now also featuring footage and references to the US coalition's "War on Terrorism" in both Iraq and Afghanistan. Just as relief organizations in our societies often uses pictures and video footage of human suffering to motivate us to give to worthy causes, these organizations do the same, playing on the emotional reactions of their audiences to find the individuals vulnerable enough to respond to their calls to action.

In the European case, where distorted Islamic ideologies are being used, often those individuals who become terrorist recruits are sensitive and care about the conflicts they learn about but are not able to read in Arabic and fall prey to teachers who tell them that they know and can interpret the Koran for them and teach them the proper Islamic response to address such suffering. This is not to say they are simple minded; quite the contrary, often it is the well-educated and sensitive individuals who read the news and care about the world and who would have been leaders if they felt they had a way of participating in their society and its political discourse. Frustrated by their lack of opportunities to help others in need and needing a meaningful role and identity, they find answers in terrorist groups.

Fourth and fifth powerful motivators are the ideas often taken from religion that one ought to sacrifice in behalf of the brotherhood of the believers. In this manner the ideas of fictive kin (Atran, 2003) and martyrdom are instilled. The individuals who respond to such calls to action are called on to depart from the frustrations of this life, reject the society that has marginalized or frustrated them, and join a group following a path that promises eternal rewards. Suddenly the individual who was previously frustrated, feeling worthless, and had little hope feels a sense of belonging, a firm identity and purpose in what will soon turn out to be a foreshortened life.

The words of a disillusioned radical (Speckhard, 2005d) living in Brussels, Belgium, explain how he was radicalized and then found his way out of it.

[3]In Brussels, we have found that there are Internet cafes where, if one logs on for a half hour or so, pop up ads appear inviting one to join the worldwide jihad.

This young man (age 24) was adopted from Rwanda by white Belgium parents. Growing up as the only black in his community, he was alienated and confused about his identity. In search for his African roots, he found Islam, converted at age 15, and started attending a radical mosque, where he fell under the influence of extremist militant wahhabist teachers. "If you say to yourself this thing is God's will, you have to do it. It's simple, if you can't read in Arabic and people tell you that you cannot understand and you have to do it. I tried very sincerely to do so. I followed everything, the prayer schedule, eating and way of drinking all in the Sunnah. But there are also ideas about jihad. At age nineteen I was ready to go to Lebanon and fight for my brother Palestinian. I didn't know politics but I had an idea we had some Muslim community, our brothers that we must defend."

Speaking about how he became ready to become a martyr, he explained,

> When I went to Morocco with my wife to her mother's house I saw Al Manar – Lebanon TV. They have a way to mix religion and politics. I can understand it because there is a true crisis in Palestine. . . . What I saw on the television was two Israeli soldiers taking big stones and breaking the bones of a Palestinian man, breaking his arm bones, his shoulders, all the bones in his hands, all the bones in his feet, his ribs, smashing them with a big rock. I'm sure they killed him or left him to die. I couldn't understand all of the Arabic (he was beginning to learn Arabic at this time) but I didn't need language to understand – it was all there in the pictures. Imagine people see that in the morning, get breakfast and see that on their television. When you see that you feel there is a unity of Muslim people. I decided to go there. I was completely crazy. I had a wife and baby but I thought I would go anyway.

Thankfully, this young man found his way out of the radical groups by studying Arabic and religion intensively until he found his own answers, independent of the militant teachers he had fallen under. Looking back at how close he came to going to be a martyr, he explained, "At that time I was a believing radical. For a radical you can die and kill for God no problem. I can die or be killed at anytime. I knew people who went to Afghanistan as bombers, the people who killed Mousad. To get the connections to go, here in Brussels, is no problem." Reflecting on how he got out, he explained, "The problem is you see all the people, politics, everything – you see through the Koran and it's your perspective." Similarly, he recalled that at that time "I was completely lost" and reflected that it was difficult to question what was being taught if you did not learn Arabic. "The difference between me [having left the radical groups] and the others is my studies."

Adopting a Terrorist Ideology as Psychological First Aid

In both motivational sets, the ideology of the terror-sponsoring group is acting as psychological first aid for the victim of other grievances. This psychological first aid is, of course, short-lived – as is its victim. Yet we see a powerful transformation take place in the human bomber who, seeing no way to change his or her circumstances, moves from a stance of a powerless victim of societal forces to becoming an actor in a worldwide or nationalistic drama that he or she has been persuaded might bring about a more just and dignified existence for those left behind. In this way, a marriage occurs between the terror-promoting ideology and the individual's psychosocial vulnerabilities emanating from traumatic and bereavement experiences in conflict zones and from marginalizing and frustrating circumstances in nonconflict zones. When this marriage occurs, all that is left is for the individual to believe that there is some significant portion of society that supports his or her stepping out onto this path (it may be the group that recruits him), which enables him to take the final steps to martyrdom. This is when societal support for suicide terrorism plays an important role in putting a martyr on the path to becoming a human bomb.

Societal Support for Suicide Terrorism

When a society deplores violence as an answer to violence and terrorism as an answer to social problems, it is unlikely that ideologies supporting suicide terrorism will resonate strongly within more than extremely limited groups of vulnerable individuals. However, when a society or significant elements of it begins to embrace an ideology in support of suicide terrorism, then these groups of potential recruits will expand exponentially. We witnessed this when a cult of martyrdom sprang up among Palestinians during the second Intifada and made their pool of potential recruits seemingly endless.

How does this occur? Again, it is a confluence of factors. When a society is so offended and even traumatized by daily living circumstances as a result of war, conflict, human rights violations, marginalization, frustrations, and daily humiliations and many individuals feel powerless to change these assaults to human dignity, the society can begin to support terrorist ideologies, especially those that espouse familiar and valued religious ideals, believing that terror acts may bring about change. In this case, the society begins to indoctrinate its children in support of suicide terror. Posters, songs, and videos may become prevalent throughout the society in support of

martyrdom (i.e., suicide terrorism), and the society as a whole may begin to resonate with the aims and ideology of the terror-sponsoring group.

It is most important that we avoid this when considering the fight against suicide terrorism. There will likely always be fringe groups that promote dying to kill; there will also always be individuals who are vulnerable to recruitment by these groups. However, we unwittingly create the circumstances in which the pool of recruits expands exponentially when we fail to address the societal factors leading to individual vulnerability and the societal support that make these groups impossible to extinguish and allows them to self-replenish faster than we can stop them. In these cases, their pool of recruits becomes so large that the terror groups can go on forever. Likewise, we must begin to address and take apart the rhetoric of terror-sponsoring organizations – addressing their ideologies by engaging with them in a discourse that can perhaps lead to more, instead of fewer, people believing that political solutions do exist and terror acts are neither necessary nor useful in bringing about a just, moral, and dignified existence.

As for the use of a hijacked version of Islam in the majority of the most current cases of suicide bombing, we must acknowledge two things. First, Islam is a religion that has always valued the struggle for three fundamental values: justice, morality, and human dignity. It is only natural that, when a terror group is hoping to motivate recruits, it appeals through Islamic traditions for recruits to take action on behalf of these fundamental values (especially if it can argue that recruits are acting in defense of self and community). Indeed, this idea of defense of self and community has been the basis of nearly all fatwas in support of martyrdom (i.e., suicide missions).

Second, a majority of the world's Islamic populations live under corrupt and autocratic regimes and face numerous human rights violations, territorial occupations, and discrimination; hence, there are many political reasons that Islamic people might gravitate to a terrorist ideology that shares with them the political goals of fighting for freedom and human dignity. Religion is simply the vehicle for uniting them and giving them the courage to fight (in whatever mistaken or brutal ways they chose) for the political goals they share. When we mistakenly believe that Islam itself is the problem and begin to assault deeply valued religious traditions and beliefs, we only fuel the fires of the terrorism ideologies currently in vogue.

There is also a contagion effect that we must consider, which occurs even with normal suicide – those who are in the family and friendship network of a suicide bomber are often so deeply affected by the act that they too begin to consider acting similarly. We found countless examples of radicalization proceeding through close friendship and relatives networks, as have other authors (Sageman, 2004). Thus, radicalization can increase geometrically once it gets going.

Suicide terrorism is a complex psychosocial and political issue and demands thoughtful, just, and carefully carried out responses to bring an end to its continued and increasing use to influence worldwide politics.

Acknowledgments

In recent years, I have had the opportunity with my colleagues to interview incarcerated suicide terrorists; potential recruits for human bombers; senders of suicide terrorists; members of terrorists groups; and, perhaps most important, family members, hostages, and close associates of accomplished bombers to ask about the bombers, the events, and the influences that motivated these individuals to choose to die in this way and thereby to construct a psychological autopsy of them. These interviews concerned Chechen, Palestinian, Lebanese, Uzbeki, and Moroccan suicide terrorism and radicalization of groups in Europe and took place in Chechnya, North Ossetia, Russia, Belarus, Uzbekistan, Lebanon, Morocco, Palestine, Israel, the Netherlands, France, United Kingdom, and Belgium. I especially thank Khapta Akhmedova, Mokhtar Benabdallaoui, Valery Krasnov, Natalia Mufel, Yoram Schweitzer, and Nadejda Tarabrina for helping me in conducting the interviews and my students Ken Reidy and Valentijn Vanrompay, who accompanied me on trips to the Palestinian territories. I am also indebted to Scott Atran, Nichole Argo, Mia Bloom, John Esposito, Rohan Gunaratna, Bruce Hoffman, Mark Juergensmeyer, Assaf Moghadam, Cerwyn Moore, Jerrold Post, Marc Sageman, Jessica Stern, Andrew Silke, Michael Taarnby, Reuven Paz, John Horgan, and many others who have contributed to this field of research and helped me refine my technique and theory and have added to my understanding of this complex phenomena through conversations, critiques of my work, and correspondence.

See also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Disasters and Mass Violence, Public, Effects of; Gulf War Syndrome, Psychological and Chemical Stressors; Nuclear Warfare, Threat of; Persian Gulf War, Stress Effects of; Posttraumatic Stress Disorder, Delayed; Vietnam Veterans, Postwar Experiences and Health

Outcomes; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Clinical.

Further Reading

Argo, N. (2003). *The banality of evil: Understanding and defusing today's human bombs*. Unpublished research proposal.

Argo, N. (2004a). *Culture, society and martyrdom*. Unpublished manuscript.

Argo, N. (2004b). *Personal communication*.

Atran, S. (2003). Genesis of suicide terrorism. *Science* **299.**

Atran, S. (2004). Mishandling suicide terrorism. *The Washington Quarterly* **27**(3), 67–90.

Bloom, M. (2005). *Dying to kill: The allure of suicide terror*. Columbia University Press.

Ganor, B. (2005). *The counter-terrorism puzzle: A guide for decision makers*. London: Transaction Publishers.

Hafez, M. (2004). *Manufacturing human bombs: Strategy, culture and conflict in the making of Palestinian suicide bombers*.

Hafez, M. (2006). *Manufacturing human bombs: The making of Palestinian suicide bombers*. United States Institute of Peace Press Books.

Jazeera, A. (Oct. 24, 2002, October). Chechens threaten to kill hostages. Retrieved September 1, 2006, from http://www.cbsnews.com/stories/2002/10/25/world/main526899.shtml

Juergensmeyer, M. (2000). *Terror in the mind of God: The global rise of religion violence*. Berkeley and Los Angeles: University of California Press.

Merari, A. (2003). *Suicide terrorism*. Unpublished manuscript.

Moghadam, A. (2003). Palestinian suicide terrorism in the second intifada: Motivations and organizational aspects. *Studies in Conflict and Terrorism* **26**(2), 65–92.

NATO Defence College. *Financial and economic aspects of the fight against terrorism, immediate report*.

Pape, R. (2005). *Dying to win: The strategic logic of suicide terrorism*. New York: Random House.

Sageman, M. (2004). *Understanding terror networks*. University of Pennsylvania Press.

Speckhard, A. (2004). Soldiers for god: A study of the suicide terrorists in the moscow hostage-taking siege. In: McTernan, O. (ed.) *The roots of terrorism: Contemporary trends and traditional analysis*. Brussels: NATO Science Series.

Speckhard, A. (2005a). Understanding suicide terrorism: Countering human bombs and their senders. In: Purcell, J. S. & Weintraub, J. D. (eds.) *Topics in terrorism: Toward a transatlantic consensus on the nature of the threat*. Washington, D.C.: Atlantic Council.

Speckhard, A. (2005b). Unpublished Beslan interviews.

Speckhard, A. (2005c). Unpublished research interviews with Belgian radicals.

Speckhard, A. (April 2005). Unpublished Palestinian militant interviews.

Speckhard, A. and Ahkmedova, K. (2005). Talking to terrorists. *Journal of Psychohistory, Fall* **33**(2).

Speckhard, A. and Ahkmedova, K. (2006). The making of a martyr: Chechen suicide terrorism. *Journal of Studies in Conflict and Terrorism* **29**(5).

Speckhard, A. and Akhmedova, K. (2004). Mechanisms of generating suicide terrorism: Trauma and bereavement as psychological vulnerabilities in human security – the Chechen case. In: Donnelly, J. (ed.). Brussels: NATO Science Series.

Speckhard, A. and Akhmedova, K. (2006a). Black widows: The Chechen female suicide terrorist. In: Schweitzer, Y. (ed.) *Female suicide terrorists*. Tel Aviv: Jaffe Center Publication.

Speckhard, A. and Akhmedova, K. (2006b). The new Chechen jihad: Militant wahhabism as a radical movement and a source of suicide terrorism in post-war Chechen society. *Democracy and Security* **2**(1), 103–155.

Speckhard, A., Tarabrina, N., Krasnov, V. and Akhmedova, K. (2004). Research note: Observations of suicidal terrorists in action. *Terrorism and Political Violence* **16**(2), 305–327.

Speckhard, A., Tarabrina, N., Krasnov, V. and Mufel, N. (2005a). Posttraumatic and acute stress responses in hostages held by suicidal terrorists in the takeover of a Moscow theater. *Traumatology* **11**(1).

Speckhard, A., Tarabrina, N., Krasnov, V. and Mufel, N. (2005b). Stockholm effects and psychological responses to captivity in hostages held by suicidal terrorists. In: Wessely, S. & Krasnov, V. (eds.) *Psychological Responses to the New Terrorism: A NATO-Russia Dialogue*. IOS Press.

Suicide, Biology of

M A Oquendo
Columbia University and New York State Psychiatric Institute, New York, NY, USA
L Giner
Fundacion Jimenez Diaz and Universidad Autonoma de Madrid. Madrid. Spain
J J Mann
Columbia University and New York State Psychiatric Institute, New York, NY, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M A Oquendo and J J Mann, volume 3, pp 538–543, © 2000, Elsevier Inc.

Introduction
A Model for Understanding Suicidal Behavior
Stressors: Triggers for Suicidal Behavior
Risk Factors Influencing the Threshold for Suicidal Behavior
Summary

Glossary

5-Hydroxyindoleacetic acid (5-HIAA)	A metabolite of serotonin.
Cerebrospinal fluid 5-hydroxyindoleacetic acid (CSF 5-HIAA)	The concentration of 5-HIAA found in the cerebrospinal fluid. The concentration of 5-HIAA in the cerebrospinal fluid is considered a reflection of serotonergic function in the central nervous system.
Major depression	A temporary or chronic mental disorder characterized by the occurrence of five of the following symptoms during a 2-week period or longer: feelings of sadness, inability to enjoy things, sleep disturbance, appetite disturbance, psychomotor changes, fatigue, difficulty concentrating, and suicidal ideas.
Model for understanding suicidal behavior	An explanation of how suicide risk factors contribute to suicidal behavior by creating a predisposition to suicidal behavior (lowering the threshold for suicidal acts) or by acting as triggers or stressors precipitating suicidal acts.
Psychosis	A temporary or chronic mental condition characterized by an inability to distinguish between reality and nonreality.
Serotonin	A neurotransmitter with multiple effects throughout the brain and with a role in a variety of psychiatric disorders and symptoms, including depression, anxiety, aggression, and suicidal behavior.
Stress	A force that disrupts the equilibrium or normal functioning of an individual's mental or physical state. Different types of stressors may precipitate suicidal behavior.
Suicide	Self-inflicted death committed by an individual with the intention to end his or her life.

This article describes how stressors interact with a low threshold for acting on suicidal ideas. The types of stress that may lead to suicidal behavior are discussed, and factors influencing the threshold for suicidal behavior are described. Stressors that contribute to suicidal behavior include negative events, social conflict, acute psychiatric conditions, acute intoxication, and contagion. Risk factors that lower the threshold for suicidal behavior include family history of suicidal behavior, low serotonergic functioning, comorbid psychiatric conditions, dysfunction of the stress-reactivity hormonal system, marital isolation, early parental loss, and a history of sexual or physical abuse. Factors that raise the threshold for suicidal behavior include religious affiliation, moral constraints, concerns about social disapproval, and feelings of responsibility toward family. An understanding of the risk factors for suicidal behavior and stressors is crucial in developing methods for the prevention of suicide attempts and completion. Because it is a combination of these psychiatric, biological, and environmental risk factors that ultimately leads to an individual's suicide, interventions that address both the role of stressors and raise the threshold for acting on suicidal impulses are critical to prevention.

Introduction

Suicide is a self-inflicted death by an individual with the intention of ending his or her life. As the eleventh leading cause of death in the United States, it accounts for approximately 30 000 deaths a year in the United States and almost 1 million deaths worldwide. In the United States, suicide is the third leading cause of death in people ages 15–24 years and the second leading cause of death in those ages 24–35 years. The World Health Organization has noted a 60% increase in suicide deaths in the past 45 years. Although the actual number of suicide attempts is more difficult to quantify, it is estimated to be 10–20 times that of completed suicides.

Psychiatric, biological, and environmental factors contribute to suicidal behavior and ultimately to suicide completion. Among environmental contributors, stress has long been recognized as having a role in the timing of suicide. However, stress may also be mediated by psychiatric or biological factors. In addition, stress alone, although important, is not a sufficient condition for suicidal behavior. Suicidal behavior also requires a vulnerability for acting on such impulses. Thus, risk factors for suicide can be understood using a model that integrates the role of stress and the threshold for acting on the resulting suicidal impulse.

A Model for Understanding Suicidal Behavior

We have described a model in which suicidal behavior is the outcome of the interaction between an individual's threshold for suicidal acts and stressors that serve as triggers for suicidal behavior. In this model, which is illustrated in Figure 1, the threshold is trait-related and refers to the propensity for suicidal behavior. In contrast, stressors, which may be considered state-related, are triggers that determine the timing and contribute to the probability of suicidal acts. Consequently, risk factors that contribute to suicidal behavior may be categorized according to whether they affect the threshold or act as stressors.

Stressors: Triggers for Suicidal Behavior

Stressors may be environmental, behavioral, biological, or a combination of these, as listed in Table 1.

Environmental Stressors

Environmental triggers can be categorized as negative events or as contagion. Negative events may include financial difficulties, interpersonal loss (death of a family member or a loved one), loss of property, conflict with parents, or other interpersonal conflict. Contagion refers to the effects of suicides in the community on an individual's decision to commit suicide. These triggers may contribute to the timing of the suicidal act.

Negative events, specifically the loss or lack of a significant relationship, have been suggested to play a key role in the deaths of (1) suicide victims without mental health care, (2) suicide victims with alcohol abuse, (3) suicide victims with personality disorders, and (4) suicide victims with mood disorders. In addition, suicide is associated with financial troubles, homelessness, and job problems. The number of adverse life events appears to play a role as well. However, stressors or negative events related to suicide may vary depending on the culture studied. Most studies published in English have been conducted in Western cultures. Similarities regarding the precipitants of suicide in Eastern and Western cultures exist, such as the number of adverse life events; the presence of severe, acute stress at the time of death; high chronic stress in the year before death; and poor quality of life in the month before death. Yet differences in terms of the type of events that may precipitate suicide have been reported. In China, for example, loss of face, described as an unexpected insult to one's reputation, is a stressor associated with suicide not generally observed in Western cultures. As such, it appears that in different cultures the sensitivity to particular environmental stressors differs.

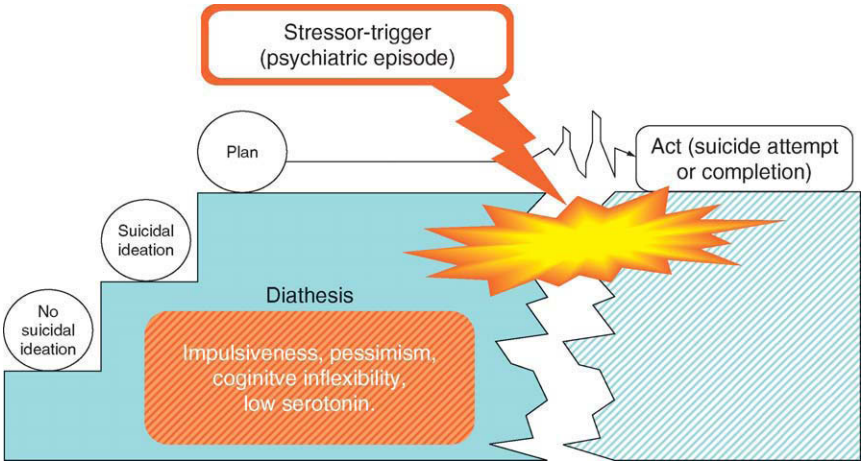


Figure 1 The diathesis–stress model for understanding suicidal behavior.

Table 1 A model for understanding suicidal behavior

	Stressors precipitating suicidal behavior	*Factors affecting the threshold for acting on suicidal ideation*
Environmental	Negative events (social, financial, or family crisis)	Marital isolation
	Contagion	Lack of religious affiliation
		Low self-esteem
		Fewer perceived responsibilities toward family
		Fewer perceived reasons for living
Behavioral	Judicial environment	Alcohol or substance abuse
	Acute substance intoxication	
Biological	Major psychiatric episode	Genetic and familial factors
	Acute substance intoxication	Low serotonergic activity (aggression)
		Stress response (childhood history of physical and sexual abuse)
		Comorbidity (alcohol or substance abuse, cluster B personality disorder, posttraumatic stress disorder)

It is possible that negative events lead to suicide at least in part by disrupting the person's social support network; this network often plays a major role in suicide prevention by providing supervision and emotional support. For example, loss of a loved one is a negative event that leads to an immediate decrease in social support, and this can trigger suicidal behavior. Similarly, social support may be acutely lowered when there is conflict with family and can thereby trigger suicidal acts. Consistent with this idea, the social networks that attempters have seem to be weaker than those of nonattempters, as measured by marital status, living alone, interactions with family members, frequency of recent moves, and the number of close friends.

Another important environmental trigger is contagion, also known as the Werther effect, which was named by Philips in 1974 after Goethe's *The Sorrows of Young Werther*. The Werther effect has been documented as a meaningful contributor to suicidal behavior in both adolescent populations and older populations. Recent outbreaks of suicide in adolescent and elderly populations have raised concerns about media coverage of suicides and the potential for precipitating more suicidal acts within communities. Presumably, individuals with a lower threshold for acting on suicidal thoughts are more likely to act on them after learning of a suicide or an assisted suicide on television or in the newspaper. Moreover, the judicial environment may affect the risk for suicidal behavior in contagion situations by changing public attitudes about the acceptability of suicide. Although there is limited data available, it is possible that the extensive media coverage of physician-assisted suicide and its legalization in one state of the United States, as well as in some European countries, has increased the rate of suicide across diagnoses, including major depression, and specifically in those who are chronically ill and depressed. In fact, physician-assisted suicides and euthanasia contribute to significant underreporting of suicides (as much as 20%) in countries such as the Netherlands, where these procedures are often not reported.

Thus, environmental factors or stressors that play a role in suicidal behavior include the disruption of social support networks, negative events, and interpersonal conflict as well as the altering of perceptions about the acceptability of suicide. These stressors are not causal, however, because the majority of individuals exposed to them do not commit suicide. Nonetheless, they may contribute to the timing and occurrence of suicidal acts in vulnerable people.

Biological and Psychiatric Stressors

Acute psychiatric illness Although depression or psychosis are not often conceptualized as stressors, it is useful to consider them as such because suicide rarely occurs in the absence of a major psychiatric disorder. The rate of mental illness in suicide victims varies from 90% in Western cultures to 60% in Eastern ones. However, even those subjects who do not meet criteria for a major psychiatric disorder seem to have some form of psychopathology. The most common mental disorders in suicide victims are depression and substance abuse, particularly alcoholism.

However, psychiatric illness alone is not sufficient to trigger suicidal behavior. Between 2 and 15% of subjects who suffer from unipolar depression ultimately commit suicide. Schizophrenia also carries a high risk of 10–15%. In bipolar disorder, the risk for suicide is generally considered to be approximately 19%; however, whereas some propose that it is nearly 50%, others question whether even 19% is an overestimate. Nonetheless, despite the high rate of suicide associated with these psychiatric conditions, most patients with these conditions do not commit suicide or even attempt suicide. Thus, the presence of a single stressor such as a psychiatric condition appears to

be a necessary, but not sufficient, cause for suicidal behavior.

Acute intoxication Substance and alcohol abuse greatly increase the risk of suicide, an effect that is especially pronounced in youth suicide. Acute substance intoxication can be considered a trigger associated with suicidal behavior. In fact, a recent review of acute alcohol use and suicidal behavior noted that 37% of completed suicides and 40% of attempted suicides involved acute alcohol use. Moreover, acute alcohol use was more prevalent among those who used violent methods of suicide such as self-immolation or gunshot.

To the extent that some individuals respond to environmental stress by using substances, acute intoxication may mediate suicidal responses to stress. The consequences of acute intoxication appear to be biological and behavioral in nature. The biological consequences of acute intoxication with alcohol include, for example, an acute increase of serotonin release, although the long-term consequence is lower serotonin function. In addition, the disinhibiting behavioral effects of intoxication may also play an important role in its association with suicidal behavior. Individuals may be more likely to act on a suicidal impulse when acutely intoxicated than when sober. Thus, acute alcohol intoxication has a complex relationship to suicidal acts.

Risk Factors Influencing the Threshold for Suicidal Behavior

Evidence for the existence of a threshold for suicidal behavior comes from two general findings. First, suicidal behavior manifests early in the course of illness, suggesting that individuals with a vulnerability for this behavior (lower threshold) are likely to express it early in the face of a stressor such as a depressive episode. Second, suicidal behavior runs in families independently of familial risk for mood disorders. Together, these factors suggest that the vulnerability to suicidal acts is not merely related to the presence of a mood disorder or to demoralization secondary to prolonged illness. Rather, suicidal behavior relates to a threshold, predisposition, or diathesis toward acting on suicidal urges. This threshold for suicidal behavior is influenced by risk factors that may be categorized as familial/genetic, biological, or environmental, as listed in Table 1.

Familial/Genetic Influences on the Threshold for Suicidal Behavior

Evidence for the presence of genetic risk factors include the presence of a family history of suicide. Although this risk factor could be thought of as facilitating suicidal behavior due to identification or contagion, it includes a genetic component as well. Adoption studies have found that, independent of the transmission of a major psychiatric disorder, adoptees who commit suicide have a sixfold higher rate of suicide among their biological relatives than do matched adoptees who do not complete suicide (None of the adopting relatives committed suicide.) Moreover, there appears to be greater concordance for suicide in monozygotic than in dizygotic twins. Other evidence for the genetic component of suicide is that the familial transmission of suicide risk to offspring appears to be independent of the transmission of psychiatric conditions. A study of Old Order Amish families has shown that family loading for mood disorders alone is not predictive of completed suicide. Rather, individuals with both a mood disorder and a family history of suicide have an increased risk for suicide.

Most genetics studies of suicidal behavior have focused on seven genes: the serotonin transporter (SERT), tryptophan hydroxylase (TPH) 1 and 2, three serotonin receptors (5-HTR1A, 5-HTR2A, and 5-HTR1B), and the monoamine oxidase promoter (MAOA), all of which are key proteins in the maintenance of serotonin function. Although there are no clear findings, the following are most promising:

1. The SERT promoter gene has two polymorphisms or variants, a short and a long form. The short form has been associated with suicidal behavior in different subpopulations, such as violent attempters and female suicide attempters. In a recent meta-analysis, suicidal behavior was reported to be more common in those with the short variant of the SERT promoter gene (OR = 1.17, CI 1.04–1.32; $p = 0.009$). Moreover, the short form has been reported to have an interaction with current life events, making those with this variant vulnerable to suicidal behavior when exposed to stress.
2. The 5-HTR1A gene encodes the 5-HT_{1A} receptor, which in depressed suicide victims has been found to have altered binding in the ventral prefrontal cortex, the region of the brain involved in decision making and restraint, among other functions.
3. The 5-HTR1B gene, which encodes the $5HT_{1B}$ receptor, appears to be related to impulsivity and aggressive behavior. This gene has two identified variants, both associated with 20% fewer receptors in the postmortem brain, although these variants were found at the same frequency in suicide and control subjects.
4. A low expressing version of the MAO gene has been reported to interact with exposure to

childhood abuse, resulting in aggression in adult males and indicating that the effect of this gene is only manifested if there are negative environmental conditions or, in other words, is dependent on a gene–environment interaction.

Thus, genetic and familial contributions to the diathesis for suicidal behavior require further inquiry but hold promise in the elucidation of suicide risk in individuals who carry the relevant variants of these genes.

Neurobiological Influences on the Threshold for Suicidal Behavior

The relationship of abnormal monoamine function to suicide risk has been studied for over 30 years. CSF 5HIAA is a metabolite of serotonin and its concentration in the cerebrospinal fluid (CSF) is considered a reflection of serotonergic function in the central nervous system (CNS). Patients with major depression who have made suicide attempts have lower levels of 5-HIAA in their CSF compared to depressed patients who do not have a history of suicide attempts. The reduction in CSF 5-HIAA appears to be more pronounced in depressed patients making planned, highly lethal attempts that more closely resemble failed suicide. Furthermore, CSF 5-HIAA correlates with characteristics of the most lethal attempt in the patient's lifetime, even when the suicide attempt is not temporally close to CSF measurements. In addition, CSF 5-HIAA seems to predict the risk of future suicide. These findings, together with those from animal studies in which the heritability and stability of CSF 5-HIAA on test–retest has been demonstrated, suggest that serotonergic dysfunction is trait-related and thus can reasonably be considered a factor that lowers the threshold for acting on suicidal thoughts.

Other measures of serotonergic function also indicate decreased serotonergic function in the CNS of suicide attempters. Fenfluramine has been widely used as a serotonin-releasing agent. Its effects on serotonin include reuptake inhibition, release from presynaptic storage granules, and possibly direct stimulation of postsynaptic receptors. We have reported that a fenfluramine challenge test elicited a lower serotonin-mediated prolactin response in patients with a major depressive episode who had a lifetime history of high-lethality suicide attempts compared with depressed patients who had a lifetime history of only low-lethality attempts. Moreover, using positron emission tomography and radioactively tagged glucose, we have reported that high-lethality suicide attempters exhibit low metabolism in the prefrontal cortex, the brain area associated with decision making and restraint. They also show impaired serotonergic responsivity. Both of these abnormalities are proportional to the lethality of their suicide attempt and are related to suicide intent and impulsivity and thus may influence lethality via these parameters. In addition, the postmortem examination of brains from suicide victims found reduced binding to the serotonin transporter sites, which was especially pronounced in the ventral and dorsal prefrontal cortex. The increased binding or number of postsynaptic 5-HT_{1A} and 5-HT_{2A} receptors in prefrontal cortex, postulated to reflect an accommodation to low serotonin, has been reported in most studies as well. Studies of brain stem, from which CNS serotonin neurons emanate, have found reduced concentrations of serotonin or 5-HIAA in suicide victims. Together, these findings suggest that serotonin dysfunction in the CNS is associated with severe and lethal suicidal behavior.

Lower serotonergic functioning may have other effects on suicidal behavior as well and may mediate the association between suicide and behaviors such as aggression and alcoholism. Among depressed individuals, those who have made suicide attempts tend to be more aggressive. Several studies show that in humans and in nonhuman primates, aggression is linked to low serotonin function. Similarly, the biological underpinnings of the predisposition to abusing alcohol, including a low serotonin function and lower CSF 5-HIAA in alcoholics during abstinence compared to immediately after the discontinuation of drinking, have been shown. Thus, individuals predisposed to alcoholism or aggression may have biological parameters, such as low central serotonergic functioning, that are also associated with suicidal behavior.

In summary, serotonergic dysfunction may influence the threshold for both attempted and completed suicide. Moreover, serotonergic function may be a mechanism whereby genetic factors influence the suicide threshold. Clearly, the relationship among genetic endowment, suicide, and serotonin functioning needs further study.

Stress Response and Suicidal Behavior

Childhood physical and sexual abuse has been associated with suicidal behavior. Those who experience corporal punishment in adolescence are reported to be at increased risk of depressive symptoms, suicidal ideation, and alcohol abuse. Suicidal ideation increases markedly with the frequency of adolescent corporal punishment for both males and females. In fact, adolescent suicide attempters are three to six times more likely than controls to have had contact with the social services department. They also commonly had notations in their chart about child abuse. In a college sample, 16% of men and 24% of women

reported sexual abuse as children. Such a history predicted depression, chronic self-destructiveness, suicidal ideation and attempts, self-mutilation, and substance abuse. The younger the person was at time of first abuse, the higher the number of suicide attempts reported. Similarly, in a sample of depressed individuals, suicide attempters were more likely to have experienced sexual abuse themselves as children, and their offspring were also at increased risk for being sexually abused and for being suicide attempters, than were depressed individuals who were not suicide attempters. These findings support the link between childhood abuse and subsequent suicidal behavior.

Childhood physical and sexual abuse has been linked to abnormalities in the stress response as mediated by the hypothalamic-pituitary-adrenal (HPA) axis. The HPA axis, with its intricate two-way interaction with the serotonergic system, has been proposed to play a key role in the correlation between early stress exposure and later impulsivity, aggression, and suicidal behavior. Indeed, suicidal behavior seems to be related to hyperactivity of the HPA axis. Among depressed individuals, those with a hyperactive HPA axis, as demonstrated by an inability to suppress cortisol in the face of external steroid administration (dexamethasone suppression test), have been reported to have a ninefold elevation in risk for suicide. Moreover, suicide victims have larger adrenal glands, which produce stress hormones and low prefrontal cortex corticotropin releasing hormone binding, consistent with a hyperactive HPA axis.

Influence of the Presence of Multiple Psychiatric Diagnoses on the Threshold

Comorbid conditions that increase the risk for attempted suicide include alcohol abuse, substance abuse, posttraumatic stress disorder, and cluster B personality disorders. In depressed inpatients, these conditions are more common among suicide attempters.

Studies have suggested that when alcoholism and depression are comorbid, suicide attempt risk is greatly increased, even when the presence of an antisocial personality disorder or other personality disorders are controlled for. Furthermore, depressed alcoholics usually report more suicidal ideation than nonalcoholic depressed patients, even when the severity of other depressive symptoms is similar. Apart from the biological pathways previously noted, other mechanisms may mediate the relationship between alcohol and suicide. For example, alcoholism may lead to unemployment, financial problems, and interpersonal problems. Psychologically, it may increase loneliness and aggression and it may inhibit coping mechanisms, all of which may lower the threshold for suicidal behavior.

Suicide attempts are also reported to be more common in depressed patients with comorbid borderline personality disorder (BPD) than in depressed patients without BPD. Moreover, subjects with major depression and comorbid borderline personality disorder are more likely to make multiple suicide attempts, and the attempts are no less medically damaging than those made by patients with major depression alone. Thus, the potential lethality of suicide attempts in patients with comorbid disorders and/or BPD should not be underestimated.

Posttraumatic stress disorder (PTSD) is frequently comorbid with depression, and when they co-occur the risk for suicidal behavior is enhanced. The relationship between PTSD and suicidal behavior appears to be mediated by the presence of cluster B personality disorder (CBPD), with both PTSD and CBPD arising as a result of earlier traumatic experiences. The assessment and treatment of comorbid conditions such as PTSD and CBPD in the context of depression may contribute to the reduction of suicide risk in this vulnerable population.

Environmental and Psychological Influences on the Threshold for Suicidal Behavior

There are environmental and psychological effects on the threshold for suicidal behavior. Environmental conditions that lower social support, such as marital isolation or living alone, are associated with suicide risk. Indeed, married women have a lower rate of completed suicide than women who are not married. In the same vein, an association between committing suicide in the context of a major affective disorder and not living with a child under the age of 18 has been reported.

However, there are clinical variables that increase the threshold for suicide attempts. These include religious or moral constraints, concerns about social disapproval, and coping and survival skills. In particular, religious affiliation seems to be associated with less suicidal behavior in depressed inpatients, and this protective effect may be mediated through moral objections to suicide and feelings of responsibility toward family. In fact, for a woman, the likelihood of completed suicide decreases as the number of children she has increases.

A variable not easily classified as influencing either threshold or trigger is low self-esteem. Although low self-esteem can be an enduring trait, it is also a symptom of depression that can be exacerbated during depressive episodes. In patients with recurrent major depressive disorder without comorbid personality disorders, those who reported feeling like a failure

were significantly more likely to have made a suicide attempt than those not reporting such feelings. In agreement with this finding, we have found that a factor of the Beck Depression Inventory that captures self-blame is associated with a history of suicidal behavior. Thus, low self-esteem may decrease the threshold for acting on suicidal thoughts in general, and its effect may be magnified during acute episodes of depression.

In summary, environmental factors related to social support and psychological attitudes, such as self-esteem or moral constraints against suicidal behavior, are important contributors to the threshold for suicide.

Summary

Risk factors can facilitate suicide by either acting as stressors or lowering the threshold for suicide. These stressors may be environmental, biological, or psychiatric in nature. Environmental stressors include negative events and contagion. Psychiatric stressors often have biological underpinnings, and acute psychiatric illness is a prominent stressor that is present in 60–90% of suicides. Acute intoxication also increases the risk of suicidal behavior by causing changes in serotonergic function and by disinhibiting the individual, making it more likely that he or she will act on suicidal impulses. These stressors act as triggers and precipitate suicidal acts.

However, whether psychiatric, environmental, or biological, stressors do not act alone in influencing suicidal behavior. Other risk factors influence suicidal behavior as well by lowering an individual's threshold for acting on suicidal ideation. Risk factors that lower the threshold for suicide can also be distinguished by three major categories: biological, psychiatric, and environmental.

Many studies of familial transmission of suicidal behavior have suggested that there is a genetic component to suicidal behavior. The strongest findings are related to the serotonin transporter gene (SERT), although there are promising results for other serotonin-related genes. Neurochemical risk factors such as lower CSF 5-HIAA levels and hyperactivity of the HPA axis are also associated with suicidal behavior and appear to be biochemical traits. These biological factors may potentially become useful predictors of suicidal behavior.

Psychiatric comorbidity may also lower the threshold for suicidal behavior. Studies suggest that in depressed patients, the presence of alcohol abuse, substance abuse, PTSD, or CBPD can increase the propensity to act on suicidal thoughts. Environmental influences that lower the threshold for suicidal behavior include marital isolation, early childhood parental loss, and a childhood history of physical or sexual abuse. An individual's threshold for suicide is also lowered by feelings of low self-esteem. Several factors appear to raise the threshold for suicidal behavior. These protective factors include religious affiliation, moral constraints, concerns about social disapproval, coping and survival skills, and family responsibility. These attitudes may act as an incentive not to commit suicide.

This article describes a comprehensive model for categorizing risk factors for suicidal behavior into those that act as stressors and those that lower the threshold for acting on suicidal ideas. This model may guide research into methods of prevention and treatment of suicidal behavior and, ultimately, save lives.

Acknowledgments

This work was supported by PHS Grant MH46745 and MH59710. We thank Ms. Ansley Roche and Sadia Chaudhury for their assistance in the preparation of this manuscript.

Further Reading

Anguelova, M., Benkelfat, C. and Turecki, G. (2003). A systematic review of association studies investigating genes coding for serotonin receptors and the serotonin transporter. II: Suicidal behavior. *Molecular Psychiatry* **8**, 646–653.

Caspi, A., Sugden, K., Moffitt, T. E., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301**, 386–389.

Cherpitel, C. J., Borges, G. L. and Wilcox, H. C. (2004). Acute alcohol use and suicidal behavior: a review of the literature. *Alcoholism, Clinical and Experimental Research* **28**, 18S–28S.

Egeland, J. A. and Sussex, J. N. (1985). Suicide and family loading for affective disorders. *Journal of the American Medical Association* **254**, 915–918.

Grunebaum, M. F., Keilp, J., Li, S., et al. (2005). Symptom components of standard depression scales and past suicidal behavior. *Journal of Affective Disorders* **87**(1): 73–82.

Linehan, M. M. and Shearin, E. N. (1998). Lethal stress: a social–behavioral model of suicidal behavior. In: Fisher, S. & Reason, J. (eds.) *Handbook of life stress, cognition and health*, pp. 265–285. New York: John Wiley & Sons.

Malone, K. M., Haas, G. L., Sweeney, J. A., et al. (1995). Major depression and the risk of attempted suicide. *Journal of Affective Disorders* **34**, 173–185.

Mann, J. J. (2003). Neurobiology of suicidal behaviour. *Nature Reviews: Neuroscience* **4**, 819–828.

Mann, J. J., Malone, K. M., Psych, M. R., et al. (1996). Attempted suicide characteristics and cerebrospinal fluid

amine metabolites in depressed inpatients. *Neuropsychopharmacology* **15**, 576–586.
Mann, J. J., Oquendo, M., Underwood, M. D., et al. (1999). The neurobiology of suicide risk: a review for the clinician. *Journal of Clinical Psychiatry* **60**(supplement 2), 7–11.
Mittendorfer-Rutz, E., Rasmussen, F. and Wasserman, D. (2004). Restricted fetal growth and adverse maternal psychosocial and socioeconomic conditions as risk factors for suicidal behaviour of offspring: a cohort study. *Lancet* **364**, 1135–1140.
Oquendo, M. A., Friend, J. M., Halberstam, B., et al. (2003). Association of comorbid posttraumatic stress disorder and major depression with greater risk for suicidal behavior. *American Journal of Psychiatry* **160**, 580–582.
Oquendo, M. A., Malone, K. M. and Mann, J. J. (1997). Suicide: risk factors and prevention in refractory major depression. *Depression and Anxiety* **5**, 202–211.
Oquendo, M. A., Placidi, G. P., Malone, K. M., et al. (2003). Positron emission tomography of regional brain metabolic responses to a serotonergic challenge and lethality of suicide attempts in major depression. *Archives of General Psychiatry* **60**, 14–22.
Phillips, M. R., Yang, G., Zhang, Y., et al. (2002). Risk factors for suicide in China: a national case-control psychological autopsy study. *Lancet* **360**, 1728–1736.

Relevant Websites

Centers for Disease Control. http://www.cdc.gov.
World Health Organization. http://www.who.int.

Suicide, Psychology of

D Lester
Richard Stockton College of New Jersey, Pomona, NJ, USA
R L Walker
University of South Carolina, Columbia, SC, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D Lester, volume 3, pp 544–548, © 2000, Elsevier Inc.

Diathesis
Stress as a Precipitant of Suicide
The Adequacy of Stressors
The Situation
Survivors
Assisted Suicide
Preventing Suicide

Glossary

Assisted suicide — Suicidal acts in which the individual is assisted. This may take the form of a prescription being written for a lethal dose of a medication or of someone else administering the lethal dose/injection.

Attempted suicide — A suicide act that the individual survives. At the lower end of medical lethality, attempted suicide may be difficult to distinguish from self-mutilation. In Europe, the trend has been to call suicidal acts that are not life-threatening parasuicide or deliberate self-harm.

Completed suicide — A suicidal act that results in the death of the individual. It is sometimes called fatal suicide.

The study of suicide is complicated by difficulties in determining intentionality, or whether the suicidal person has engaged in intentional self-harm. Recently, suicide nomenclature has evolved such that clinicians, researchers, and others can now use a common language to understand the suicide-related behavior. Beyond assisted suicide, attempted suicide, and completed suicide, suicidal ideation, defined as any self-reported thoughts of engaging in suicide-related behavior, is also commonly used. In addition, parasuicide is used rather than attempted suicide to describe self-destructive behavior in the absence of clear intent of suicide. The term suicidality should be avoided because it refers to a broad range of suicide-related behaviors.

The 2001 National Strategy for Suicide Prevention noted that 650 000 people receive emergency care for suicide attempts and 30 000 die by suicide each year in the United States. Given the loss of life and impact on families and communities, suicide has been declared a serious public health problem. The goals of the National Strategy are to (1) promote awareness of suicide as a public health problem; (2) develop broad-based support for suicide prevention; (3) develop and implement strategies to reduce the stigma associated with mental illness, substance abuse, and suicide; (4) develop and implement community-based suicide prevention programs; (5) promote efforts to reduce access to lethal methods of self-harm; (6) implement

effective training to assess at-risk individuals and provide effective intervention; (7) develop clinical and professional practices; (8) increase access to mental health services; (9) improve the reporting and media portrayal of suicide-related behavior; (10) promote research on suicide and suicide prevention; and (11) improve and expand surveillance systems. The overarching objective is a comprehensive and integrated approach to reducing the loss associated with suicide and related behaviors. Many other nations have also developed national strategies for reducing the incidence of suicide.

The 2002 Institute of Medicine (IOM) report for the United States is an excellent and comprehensive resource for understanding suicide's history; impact; manifestations across gender, ethnic, age, and other demographic groups; and protective factors or suicide buffers. For example, although African Americans were previously thought to be protected from suicide, studies show that in the 1980s and 1990s Black men ages 25–34 years demonstrated suicide rates comparable to that of White men in the same age group. The most recent data and trends are cited in the IOM report, including the cost to society due to medical expenses for emergency intervention and treatment, lost productivity, and reduced wages. The value of lost productivity alone was estimated to be $11.8 billion in 1998.

Suicide can be studied at the societal or at the individual level. Sociological research has focused primarily on the societal rate of completed suicide, a rate that is relatively stable over time. At the individual level, in contrast, both completed and attempted suicides have been studied and both behaviors are difficult to predict. Suicide is rare, with the completed suicide rate, for example, seldom rising above 40 per 100 000 per year in any nation. The rarity of suicidal behavior means that prediction instruments (typically based on personal data and psychiatric and psychological measures) will have a high rate of false positives (people predicted to be suicides who, in fact, will not kill themselves). Differences between the predictability of suicide at societal and individual levels has an analogy in radioactive decay, in which the half-life of a mass of, say, uranium is easily determined and constant, whereas predicting which particular uranium atom will decay next is impossible.

Suicide fits well with the diathesis–stress model of psychiatric disorder, in which only those individuals with a predisposition (diathesis) to develop a disorder do so when faced with particular stressors. It should be noted, however, that both factors can influence one another – a diathesis can make it more likely that a person will encounter stressors and a stressor can strengthen a diathesis. For example, a stress such as the loss of a loved one can be more stressful for a person who has few friends than for someone without that diathesis. Likewise, a tendency to have few friends may be increased by the experience of interpersonal losses.

Diathesis

Many predisposing factors for suicide have been identified by research.

1. Neurophysiological dysfunction, especially problems in serotonergic pathways, may underlie the psychiatric disorders for which the risk of suicide is high.
2. The majority of suicides are found to have been suffering from a psychiatric disorder. In particular, suicide is common in those suffering from affective disorders such as a major depressive disorder or a bipolar affective disorder (manic–depression), substance abuse (both drug and alcohol abuse), and schizophrenia.
3. Childhood experiences such as the loss of a parent (or other primary caregiver) through death, especially if the death is from suicide, or divorce and physical or sexual abuse are commonly found in the personal histories of suicides. Lester studied 30 suicides famous enough to have a biography written about them (such as Ernest Hemingway) and found that exactly half had experienced the loss of a parent or other loved one. Fourteen of the 15 suicides who had experienced such a loss had experienced this loss between the ages of 6 and 16. However, it is not clear whether such experiences increase the risk of suicide directly or whether they increase the risk of psychiatric disorders, which in turn increases the risk of suicide. For example, a history of physical and sexual abuse is associated with a wide spectrum of later psychiatric disorders, including depression, anxiety, eating disorders, and substance abuse, in addition to suicidal behavior.
4. The mood of potential suicides is characterized by depression. Suicidal individuals obtain high scores on self-report measures of depression, and the level of depression of those attempting suicide increases with the seriousness of their suicidal intent. Depression includes many components: physical symptoms (such as sleep and appetite problems), mood symptoms (such as depression and guilt), and cognitive symptoms (such as hopelessness). Suicidal individuals tend to be hopeless, believing that things will not improve and will probably worsen in the future. Aaron Beck has

found that self-report measures of hopelessness are stronger predictors of suicidal intent than self-report measures of depression. Edwin Shneidman has called the psychological pain experienced by potential suicides psychache.

5. Suicidal individuals appear to think differently from nonsuicidal individuals. Research has found that suicidal individuals have a reduced ability to generate alternatives other than death for the problems they are facing; they tend to think in dichotomous categories, polarizing their evaluations into extremes, such as good versus bad or right versus wrong; and they tend to think rigidly.
6. In addition, the risk of suicide is strongly associated with simple demographic characteristics. Suicide rates in the United States are higher in men than in women, in the elderly than in the young, in Whites than in African Americans, and in the divorced and widowed than in the single and married. These demographic correlates of suicide risk vary from country to country, however.

In 2005, Joiner developed a comprehensive interpersonal-psychological theory of attempted and completed suicide and hypothesized that suicidal people must (1) be capable of lethal self-injury, (2) perceive that they are a burden to loved ones, and (3) believe that they are not connected to a valued group or relationship. These three necessary precursors answer the question of why people do or do not die by suicide. For example, few people die by suicide, relatively speaking, and interpersonal-psychological theory suggests that the acquired ability to inflict self-harm requires courage, fearlessness, and habituation to self-harm. Joiner asserted that individuals must also see themselves as ineffective and burdensome to loved ones. In contrast, a sense of belonging to a valued group acts as a suicide buffer. In sum, the confluence of three factors must be in place. However, those who engage in repeated attempts at suicide are at high risk for suicide.

Joiner's theory explains several research findings. For example, the relatively low rate of suicide despite the high lifetime prevalence of lifetime suicide thoughts is explained by the absence of two or more factors (such as the incapability of self-injury and a sense of connectedness). Thus, not everyone who considers suicide or feels hopeless will attempt suicide. Note that those who die by suicide use more lethal means and therefore demonstrate a stronger propensity to self-harm. Although women are less likely to die by suicide than men, those women who die are as likely as men to use very lethal methods (e.g., a gun and hanging). The role of neurophysiological factors in suicide is related to the impact of neurotransmitters on negative emotionality, which may, in turn, incite feelings of burdensomeness and contributes to a lesser feeling of belongingness.

Stress as a Precipitant of Suicide

Suicides are found to have experienced higher levels of stressful events in the past year than nonsuicidal individuals. Not only is the level of stressful events higher, but typically there is a further increase in this already high level in the weeks immediately prior to the suicide. These conclusions are based on retrospective studies of the lives of completed suicides in the hours, days, weeks, and months prior to their death – a procedure called a psychological autopsy. In a psychological autopsy, mental health professionals, usually guided by an interview schedule, seek out and interview friends, colleagues, and relatives of the suicide and inquire into the suicide's recent behaviors and psychological state.

Commentators have suggested that men become suicidal more as a reaction to performance failure, whereas women become suicidal more as a reaction to romantic conflict. However, research has indicated few differences between the stressors encountered by female and male suicides. Large differences in precipitating stressors by age have been observed. Stressors experienced by adolescent and young adult suicides tend more often to involve negative interpersonal events, such as friction and break-ups with loved ones and problems involving work, finances, and the criminal justice system. In contrast, stressors of elderly suicides tend more often to involve medical problems and intrapsychic states such as depression. These differences by age are found both in psychological autopsy studies and in content analyses of suicide notes.

Similar differences have been observed in attempted suicides and those thinking about the possibility of killing themselves, who also are found to have experienced more recent stressful events than nonsuicidal psychiatric patients or general medical patients. Two studies from the 1990s illustrate the role of stress. Paul Dean and colleagues found that suicidal ideation in college students was positively associated with self-ratings of negative life events in the past year and, in addition, depression, hopelessness, and perfectionism, and was negatively associated with scores on a reasons-for-living scale. John Reich and colleagues found that suicidal ideation in a sample of elderly adults (mean age 70 years) was associated with worsening health in the prior 6 months and, in addition, self-reported feelings of helplessness, fatalism, confused thinking, and low self-esteem.

The Adequacy of Stressors

Older adults who commit suicide are typically seen as having experienced more adequate stressors than younger adults and adolescents who kill themselves. Elderly suicides are frequently suffering from medical problems, often severe, and have experienced a succession of losses (of loved ones and friends, of employment and work, and of physical abilities). Thus, the precipitants of their suicides appear to involve severe stress. Most of the examples of rational suicide involve such cases, and indeed assisted-suicide laws passed in the United States and other nations are usually restricted to those who are terminally ill. In contrast, the precipitants of suicide in adolescents and young adults are sometimes seen as inadequate causes of suicide. Such suicides are, therefore, viewed as impulsive and not rational and are less predictable.

The Situation

The situation in which potential suicides find themselves may also play a role in precipitating their suicides. The easy availability of lethal methods for suicide may play a role. Research on the effects of restricting easy access to methods for suicide (such as detoxifying domestic gas, placing emission controls on cars, and restricting the number of pills prescribed for psychiatric disorders) has shown that this can be an important tactic for preventing suicide.

The publicity given to suicides, especially suicides by celebrities, has been shown to lead to an increase in the suicide rate in the following few days, especially among those of the same age and sex as the celebrity. Suicide among peers can have a contagion effect, precipitating suicidal behavior in those with those characteristics that predispose the individual to suicide. Suicides occurring in schools have aroused particular concern, and there are now guidelines for staff in responding to the suicide of a student.

Survivors

The experience of having a loved one complete suicide can be extremely traumatic for those left behind. In addition to the normal grieving process, there is additional stress from feelings of guilt and responsibility (could something have been done to help the suicide and prevent their death?), anger (over their suicide), and stigma from others who may view the survivors with suspicion or simply not know how to respond.

In addition, some suicides kill themselves in such a way that their loved ones discover the body. The discovery of the body of a loved one drowned, hanging from a hook, with cut wrists, or, even worse, grossly disfigured by a bullet is extremely traumatic. It creates a final picture of the loved one that is difficult to exorcize.

Survivors have become very active in their communities and in professional organizations devoted to suicide prevention. They have formed self-help groups, sometimes with the involvement of clinicians, to meet and assist one another in dealing with the complicated grief that accompanies the suicide of a loved one. Books have been written for survivors and the membership of the American Association of Suicidology has a large contingent of survivors.

Assisted Suicide

The reluctance of people to kill themselves with violent methods and the desire of suicidal individuals to have others assist them in their suicides have led to widespread debate about assisted suicide. Derek Humphry published the book *Final Exit*, which details the medications and dosages for suicide – a do-it-yourself book that became a best seller in the United States. Similar books in other nations (such as France) have been banned.

In physician-assisted suicide, typically the goal is to have physicians provide a sufficient amount of a lethal medication for the client to complete suicide. In most regions that have legalized this process (e.g., in Oregon and the Netherlands), this process is restricted to those who are terminally ill, and there is a detailed procedure for those applying for physician-assisted suicide that must be followed.

However, on occasion, physicians and relatives have assisted the suicidal individual in the actual killing, for example, by administering the lethal injections or by shooting the person. This is illegal in almost all nations and jurisdictions, and those assisting in the actual killing can be, and indeed sometimes are, charged with murder. Nevertheless, such incidents have occurred and continue to occur. The line between assisting someone to commit suicide and murdering someone is, therefore, blurred in these cases, and society is vigorously debating these issues at the present time.

Preventing Suicide

The prevention of suicide uses several strategies. Many communities now have suicide prevention centers that maintain telephone counseling services staffed primarily by well-trained lay people who can provide crisis counseling for those who are suicidal. The American Association of Suicidology has a directory of such centers in the United States and also

inspects and certifies the centers if asked to do so. Befrienders International, Life Line, and the International Federation of Telephone Emergency Services are three international organizations that have established and coordinated crisis counseling and suicide prevention services in other nations.

Educational programs have been established to inform students in school about suicide, especially about recognizing signs of suicidal intent in their peers, and to provide them with resources to which they can refer their suicidal peers. Educational programs have also occasionally been established for general practitioners and family physicians so that they can diagnose depression in their patients more accurately and prescribe antidepressants more effectively.

Clinicians have been developing counseling techniques especially designed for suicidal clients to supplement the general systems of psychotherapy (such as transactional analysis and reality therapy) that guide therapists. These new techniques have been primarily in the field of cognitive therapy and are based on the results of research on the cognitive processes of suicidal individuals discussed earlier.

The 2002 report from the IOM examined prevention and medical and psychosocial intervention, but also looked at barriers to research, treatment, and effective prevention. For example, suicide is more likely to occur within 1 month after discharge from psychiatric institutions than at other times. The report noted that a significant gap in clinical trials is the absence of high-risk patients. Because effective treatments (psychopharmacological and psychosocial) involve long-term maintenance in which the individual is required to have contact with a health-care professional, adequate health-care insurance coverage is essential. School-based programs that target high-risk school-age children are essential for suicide prevention among youth and seem to result in increased social support, self-efficacy, and self-esteem among youth at risk for suicide, but research on these intervention programs is complicated by poor control groups and limited long-term follow up.

See Also the Following Articles

Depression and Manic-Depressive Illness; Suicide, Biology of; Suicide, Sociology of; Survivor Guilt.

Further Reading

Clark, R. V. and Lester, D. (1989). *Suicide: closing the exits*. New York: Springer-Verlag.

Dean, P. J., Range, L. M. and Goggin, W. C. (1996). The escape theory of suicide in college students. *Suicide and Life-Threatening Behavior* **26**, 181–186.

Freeman, A. and Reinecke, M. A. (1993). *Cognitive therapy of suicidal behavior.* New York: Springer.

Henry, A. F. and Short, J. F. (1954). *Suicide and homicide*. New York: Free Press.

Humphry, D. (1991). *Final exit*. Eugene, OR: The Hemlock Society.

Institute of Medicine of the National Academies. (2002). *Reducing suicide: a national imperative*. Goldsmith, S. K., Pellmar, T. & Bunney, W. E. (eds.). Washington, DC: The National Academies Press.

Joiner, T. (2005). *Why people die by suicide*. Cambridge, MA: Harvard University Press.

Leenaars, A. A. and Wenckstern, S. (eds.) (1991). *Suicide prevention in schools*. New York: Hemisphere.

Lester, D. (1997). *Making sense of suicide*. Philadelphia: Charles Press.

Lukas, C. and Seiden, H. M. (1987). *Silent grief*. New York: Scribners.

O'Carroll, P. W., Berman, A. L., Maris, R. W., et al. (1996). Beyond the Tower of Babel: a nomenclature for suicidology. *Suicide and Life-Threatening Behavior* **26**, 237–252.

Reich, J. W., Newson, J. T. and Zaura, A. J. (1996). Health downturns and predictors of suicidal ideation. *Suicide and Life-Threatening Behavior* **26**, 282–291.

Shneidman, E. S. (1996). *The suicidal mind*. New York: Oxford University Press.

U.S. Department of Health and Human Services (2001). *National strategy for suicide prevention: goals and objectives for action*. Rockville, MD: Public Health Service.

Suicide, Sociology of

D Lester
Richard Stockton College of New Jersey, Pomona, NJ, USA
R M Fernquist
Central Missouri State University, Warrensburg, MO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by D Lester, volume 3, pp 549–552, © 2000, Elsevier Inc.

Durkheim's Theory
Henry and Short's Theory
Naroll's Theory
Culture Conflict
Measuring Regional Stress
Multivariate Studies of Regional Suicide Rates
Nonfatal Suicide Behavior

Glossary

Attempted suicide	A suicide act that the individual survives. Attempted suicide is at the lower end of medical lethality and may be difficult to distinguish from self-mutilation. In Europe, the trend has been to call suicidal acts that are not life-threatening parasuicide or deliberate self-harm.
Completed suicide	A suicidal act that results in the death of the individual. It is sometimes called fatal suicide.

Sociological research into suicide has focused primarily on the rate of completed suicide in regions (nations or regions within nations), a rate that is relatively stable over time. Ecological studies examine suicide rates of a set of regions at one point in time, usually to see which social characteristics of the regions predict the suicide rates. Time-series studies examine suicide rates of one region each year for a period of time, usually to see which social characteristics predict the yearly suicide rate. Time-series cross-sectional studies examine suicide rates of a number of different countries simultaneously for a period of time. Here multivariate ecological studies of suicide are reviewed. Three major theories of suicide are presented: Emile Durkheim's, Andrew Henry and James Short's, and Raoul Naroll's. The role of culture conflict in increasing the rate of suicide is illustrated. Finally, the possibility of sociological studies of nonfatal suicide is discussed.

Durkheim's Theory

The first major theory of the social suicide rate was proposed in 1897 by Emile Durkheim. Durkheim argued that two broad social characteristics are responsible for determining the social suicide rate: social integration and social regulation. Social integration is the extent to which the members of the society are bound together in social networks. Very high levels of social integration lead to high rates of altruistic suicide, whereas very low levels lead to high rates of egoistic suicide. Social regulation is the extent to which the desires and behavior of the members of the society are restricted by social norms and customs. Very high levels of social regulation lead to high rates of fatalistic suicide, whereas very low levels lead to high rates of anomic suicide.

There has been a great deal of research purporting to test this theory, for example, exploring the association between divorce rates and suicide. Divorce is presumed to weaken the extent of social integration and reflect weak social regulation, and therefore higher divorce rates should be positively associated with higher suicide rates. Both ecological studies and time-series studies have confirmed this predicted positive association between divorce rates and suicide rates.

However, research rarely examines the roles of both social integration and social regulation separately, partly because it is difficult to operationalize the two social characteristics independently of one another. The few studies that have attempted to do so have found that measures of social integration are stronger correlates of suicide rates than measures of social regulation.

Furthermore, no research has attempted to classify suicides in the societies under study into Durkheim's four types. Durkheim proposed, for example, that low levels of social integration result in high levels of egoistic suicide. Therefore, a proper test of Durkheim's theory requires the dependent variable to be the rate of egoistic suicide. The rate of egoistic suicide is not identical to the total social suicide rate, which is, instead, the combined rate of all four of Durkheim's types of suicides.

Henry and Short's Theory

Henry and Short proposed a theory based on both psychoanalytic theory and the frustration–aggression hypothesis. They assumed that the primary target of aggression for a frustrated individual is another person. What inhibits this other-oriented aggression and results in the aggression being directed inward onto the self?

Henry and Short proposed that the primary factor was the strength of external restraints on behavior. When behavior is required to conform rigidly to the demands and expectations of others, the role of others in the responsibility for the self's frustration and misery is strong. As a consequence, other-oriented aggression is legitimized. When external restraints are weak, the self must bear the responsibility for the frustration and misery. Thus, other-oriented aggression is not legitimized, and self-directed aggression becomes more likely.

These proposals lead to interesting predictions, many of which have been confirmed. The oppressed in a society have clear external sources to blame for their misery – their oppressors. Therefore, other-oriented aggression is legitimized for them, and they will tend to have higher rates of assault and, in the extreme, homicide. In contrast, the oppressors in the society have fewer external sources to blame for their misery because, as the dominant group, they have tremendous opportunities for advancement and gratification. Therefore, the oppressors in the society will tend to have higher rates of depression and, in the extreme, suicide. These predicted differences in the suicide and homicide rates are found for African Americans and Whites in the United States.

Henry and Short's theory can also explain the positive association between the quality of life and suicide rates. When the quality of life in a nation is high, there are few external sources to blame for one's misery, and so suicide will tend to be more common. This association has been confirmed in ecological studies both of nations and of U.S. states.

Naroll's Theory

Raoul Naroll, an anthropologist, proposed a theory of social suicide rates that has relevance to stress. Naroll proposed that suicide was more likely in those who were socially disoriented, that is, in those who lack or lose basic social ties (such as those who are single or divorced). Because not all socially disoriented people commit suicide, there must be a psychological factor that makes suicide a more likely choice when a person is socially disoriented, and Naroll suggested that it was the person's reaction to thwarting disorientation contexts. Such contexts involve a weakening of the person's social ties as a result of the actions of other people or him- or herself (but not as a result of impersonal, natural, or cultural events). Being divorced by a spouse or murdering one's spouse are examples of such contexts, but storm damage to property or losing a spouse to cancer are not. In thwarting disorientation contexts, some people commit protest suicide, which Naroll defined as voluntary suicide committed in such a way as to come to public notice (and, therefore, excludes unconsciously motivated suicides and culturally motivated suicides such as suttee).

Durkheim's theory of suicide refers more to steady-state characteristics of the society as major explanatory variables, whereas Naroll's theory suggests the role of sudden, acute changes in a society – societal stressors. Furthermore, Narroll's theory is phrased in a way that makes it more usefully applied to individuals as well as to societies as a whole.

Culture Conflict

Cultures often come into conflict, a source of stress that may increase the suicide rate. For example, the conflict between traditional Native American culture and the dominant Anglo-American culture has often been viewed as playing a major role in precipitating Native American suicide. Philip May reviewed three hypotheses that might have relevance here. In the social disorganization hypothesis, the dominance of the Anglo-American culture is viewed as forcing Native American culture to change and as eroding traditional cultural systems and values. This changes the level of social regulation and social integration, important causal factors for suicide in Durkheim's theory of suicide.

A second hypothesis focuses on cultural conflict. The pressure from the educational system and mass media on Native Americans, especially on youth, to acculturate, a pressure that is opposed by their elders, leads to great stress for the youth. A third hypothesis focuses on the breakdown of the family in Native American tribes. Parents are often unemployed, substance abusers, and in trouble with the law, and divorce and desertion of the family by one or both parents are common. The suicide rates of three groups of Native Americans in New Mexico, the Apache, Pueblo, and Navajo (whose suicide rates were 43.3, 27.8, and 12.0 per 100 000 per year, respectively), were in line with their levels of acculturation as rated by May (high, moderate, and low, respectively).

Measuring Regional Stress

Linsky and co-workers measured the stress level of each of the U.S. states in three areas. For economic stress, they used business failures, unemployment claims, strikes, bankruptcies, and mortgage foreclosures. For family stress, they used divorces, abortions, illegitimate births, and fetal and infant deaths. For community stress, they used disasters, new housing starts, new welfare recipients, high school dropouts, and interstate migration. They found that states with

higher levels of stress also had higher suicide rates. Interestingly, however, the ratings of the stress level of each state based on responses from a national survey, which asked residents about their perceived level of stress, were not associated with the state suicide rates.

Multivariate Studies of Regional Suicide Rates

The accuracy of official national and regional rates of suicide has been questioned by Jack Douglas and others, who have argued that these rates are biased by the values of the local coroners and medical examiners and of the resident populations. However, despite the fact that regions do have different standards for classifying a death as suicide, Peter Sainsbury and Brian Barraclough showed that the suicide rates of immigrants to one country from different nations are in almost the same rank order as the suicide rates of nations of origin.

The suicide rates of 71 nations reporting to the World Health Organization are shown in Table 1 for males and females. An inspection of the table reveals two basic findings: (1) in all countries, males have higher suicide rates than females, which suggests that differential gender socialization (e.g., males are socialized to be more aggressive and to keep their feelings inside) significantly impacts suicide rates; and (2) suicide rates vary significantly among the 71 countries, which suggests that cultures differ in the degree to which suicide is viewed as an acceptable behavior.

Table 1 Suicide rates around the world[a]

Country	*Year reported*	*Males*	*Females*	*Country*	*Year reported*	*Males*	*Females*
Lithuania	2002	80.7	13.1	Norway	2001	18.4	6.0
Russian Federation[b]	2002	69.3	11.9	Canada	2000	18.4	5.2
Belarus	2001	60.3	9.3	Chile	2001	18.2	3.0
Kazakhstan	2002	50.2	8.8	United States	2000	17.1	4.0
Latvia	2002	48.4	11.8	Hong Kong	2000	16.1	10.1
Estonia	2002	47.7	9.8	Northern Ireland	2002	15.9	3.5
Ukraine	2002	46.7	8.4	Puerto Rico	2000	15.2	1.4
Hungary	2002	45.5	12.2	Turkmenistan	1998	13.8	3.5
Slovenia	2002	44.4	10.5	Thailand	2000	13.5	3.7
Japan	2002	35.2	12.8	Argentina	2001	13.4	3.5
Finland	2002	32.3	10.2	Netherlands	2003	12.7	5.9
Belgium	1997	31.2	11.4	Spain	2001	12.2	3.7
Austria	2002	30.5	8.7	Uzbekistan	2000	11.8	3.8
Croatia	2002	30.2	10.0	El Salvador	1999	11.6	5.4
Uruguay	2000	29.0	5.5	Costa Rica	2002	11.6	2.0
France	2000	27.9	9.5	Singapore	2001	11.5	6.9
Moldova, Republic of[c]	2002	27.9	5.2	Italy	2001	11.1	3.3
Switzerland	2000	27.8	10.8	England and Wales	2002	9.8	2.8
Poland	2002	26.6	5.0	Israel	1999	9.8	2.3
Bulgaria	2002	25.6	8.3	Venezuela	2000	8.8	1.5
Korea, Republic of	2002	24.7	11.2	Panama	2000	8.4	1.3
Czech Republic	2002	24.5	6.1	Columbia	1999	8.1	2.4
Romania	2002	23.9	4.7	Bahrain	2000	7.2	0.3
Slovakia	2000	22.6	4.9	Brazil	2000	6.4	1.6
Trinidad & Tobago	1998	22.2	4.7	Mexico	2001	6.3	1.3
Cuba	2001	21.4	8.0	Ecuador	2000	6.0	2.6
Denmark	1999	21.4	7.4	Albania	2001	5.5	2.3
Ireland	2001	21.4	4.1	Greece	2001	5.3	0.9
Germany	2001	20.4	7.0	Armenia	2002	4.0	0.7
Australia	2001	20.1	5.3	Paraguay	2000	3.9	1.7
New Zealand	2000	19.8	4.2	Georgia	2001	3.4	1.1
Scotland	2002	19.7	5.9	Guatemala	1999	3.4	0.8
Kyrgyzstan	2002	19.1	4.0	Philippines	1998	1.8	0.6
Sweden	2001	18.9	8.1	Azerbaijan	2002	1.8	0.5
Portugal	2002	18.9	4.9	Egypt	2000	0.1	0.0
Mauritius	2000	18.8	5.2				

[a]Rate per 100 000 population from latest year reported, ranked according to male suicide rates.
[b]Excluding Chechnya.
[c]Excluding Transdinestria.
Source: World Health Organization Statistical Information Systems website.

Cutright and Fernquist term this phenomenon the culture of suicide.

Simpson and Conklin identified two clusters of variables that were associated with national suicide rates: (1) a cluster that had the highest loading from the percentage of Muslims in the population and (2) a cluster that seemed to assess economic development. Suicide rates were lower in nations with less economic development and where Islam was a more important religion. Two other clusters (Christianity and the Eastern bloc) were not associated with suicide rates.

In similar study of the social correlates of national suicide rates, Lester identified 13 orthogonal (independent) factors for social variables, only one of which was associated with suicide rates, a factor that seemed to measure economic development (with high loadings from such social variables as low population growth and high gross domestic product per capita). Thus, these two studies agreed in finding that economic development is associated with higher suicide rates. Cutright and Fernquist reported that classical Durkheimian variables measuring social integration (i.e., divorce, fertility, and religiosity), as well as female labor-force participation, were strong predictors of cross-national suicide rates for both males and females of different age groups.

For the United States, a similarly designed study conducted by Lester identified a cluster of variables that seemed to measure social disintegration (high divorce and interstate migration rates, low church attendance, and high per capita alcohol consumption), and this was the strongest correlate of the suicide rates of the states.

Nonfatal Suicide Behavior

Almost all sociological research and theories have hitherto been proposed for fatal (completed) suicide. Nonfatal suicidal behavior has been relatively ignored. It is important that better epidemiological studies of attempted suicide and suicidal ideation be conducted so that sociologists can develop theories of these behaviors. The World Health Organization has sponsored a comparative epidemiological study of attempted suicide in 15 sites. Unfortunately, the sites are limited to cities in European nations and do not encompass these nations as a whole. However, the study is innovative and may lay the groundwork for more comprehensive epidemiological studies in the future. Linsky and colleagues have obtained estimates of the incidence in the previous year of suicidal ideation in U.S. states and sought social correlates of this. Thus, such research is possible, and it is to be hoped that more studies are planned along these lines.

See Also the Following Articles

Affective Disorders; Aggression; Cultural Transition; Depression Models; Suicide, Biology of; Suicide, Psychology of.

Further Reading

Cutright, P. and Fernquist, R. M. (2000). Effects of societal integration, period, region, and culture of suicide on male age-specific suicide rates: 20 developed countries, 1955–1989. *Social Science Research* **29**, 148–172.

Cutright, P. and Fernquist, R. M. (2000). Societal integration, culture, and period: their impact on female age-specific suicide rates in 20 developed countries, 1955–1989. *Sociological Focus* **33**, 299–319.

Cutright, P. and Fernquist, R. M. (2003). The gender gap in suicide rates: an analysis of twenty developed countries, 1955–1994. *Archives of Suicide Research* **7**, 323–339.

Douglas, J. D. (1967). *The social meanings of suicide*. Princeton, NJ: Princeton University Press.

Durkheim, E. (1897). *Le suicide*. Paris: Felix Alcan. [English edition: (1951). *Suicide*. New York: Free Press.]

Henry, A. F. and Short, J. F. (1954). *Suicide and homicide*. New York: Free Press.

Lester, D. (1988). *Why women kill themselves*. Springfield, IL: Charles Thomas.

Lester, D. (1989). *Suicide from a sociological perspective*. Springfield, IL: Charles Thomas.

Lester, D. (1994). *Patterns of suicide and homicide in America*. Commack, NY: Nova Science.

Lester, D. (1994). *Patterns of suicide and homicide in the world*. Commack, NY: Nova Science.

Lester, D. (1995). Thwarting disorientation and suicide. *Cross-Cultural Research* **29**, 14–26.

Linehan, M. (1973). Suicide and attempted suicide. *Perceptual and Motor Skills* **37**, 31–34.

Linsky, A. S., Bachman, R. and Straus, M. A. (1995). *Stress, culture, and aggression*. New Haven, CT: Yale University Press.

Sainsbury, P. and Barraclough, B. M. (1968). Differences between suicide rates. *Nature* **220**, 1252.

Simpson, M. E. and Conklin, G. H. (1989). Socioeconomic development, suicide and religion. *Social Forces* **67**, 945–964.

Van Winkle, N. W. and May, P. A. (1986). Native American suicide in New Mexico, 1959–1979. *Human Organization* **45**, 296–309.

Relevant Website

World Health Organization Statistical Information Systems website. http://www3.who.int/whosis/menu.cfm.

Surgery and Stress

C Vögele
Roehampton University, London, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C Vögele, volume 3, pp 553–554, © 2000, Elsevier Inc.

Stressful Characteristics of the Surgical Situation
Responses of Surgical Patients at Emotional, Cognitive, and Physiological Levels
Surgical Stress and Postoperative Recovery
Psychological Preparation for Surgery

Glossary

Psychological preparation for surgery	Psychological methods aimed at reducing preoperative anxiety and aiding the recovery process.
Surgical outcomes	Variables reflecting the quality of postoperative recovery, for example, postoperative pain, pain medication length of stay, clinical recovery measures (e.g., wound healing), physiological indices (heart rate, blood pressure, palmar sweat gland activity, and cortisol level), negative affect, and satisfaction.
Surgery	A subset of invasive medical procedures involving some form of anesthesia.

The introduction into the health-care system and hospitalization can be a stressful experience on their own. Among the aspects of hospitalization considered to be most stressful, the following were identified in a sample of elderly medical inpatients: communication problems with health-care professionals; diagnostic and therapeutic procedures; the hospital environment (noise, rigid routines, etc.); worries about the home situation and the separation from home; insufficient information about diagnosis and prognosis; and fear of dependency, loss of autonomy and control. In addition to uncertainties about their illnesses (because of unclear communication) and the unfamiliar surroundings of a hospital ward, therefore, patients encounter additional stress because they must undergo medical procedures and examinations. Even some outpatient procedures such as dental treatments or blood donation can be stressful. Normally patients report increases in anxiety in anticipation of procedures such as surgery, endoscopy, cardiac catheterization, cancer screening, and chemotherapy. In addition to the anticipatory anxiety associated with the procedure itself, diagnostic procedures such as cancer screening are characterized by a prolonged period of anxiety between screening and receiving the result.

Of all medical procedures, surgery is perhaps the most threatening event because it contains many unpredictable and uncontrollable features such as losing consciousness due to the administration of a general anesthetic, the anticipation of postoperative pain, and the surgical trauma related to the incision. This article concentrates on surgical procedures and describes the stressful characteristics associated with being hospitalized for surgery. We review the literature on the psychological and physiological responses to the experience of surgery and summarize findings from studies investigating psychological preparation for surgery.

Stressful Characteristics of the Surgical Situation

There are obvious potential threats for the surgical patient: anesthesia, pain, physical restriction, life-threatening procedures, and being away from home. Evidence suggests that the lack of predictability and control are significant contributors to the stressful experience of surgical patients. It is usually the case that we need to be able to predict an event in order to be able to control it. But the reverse does not follow – being able to predict an event does not necessarily mean we can control it. Such is the case for elective surgery, which accounts for the vast majority of surgical interventions.

This evidence confirms that there are common characteristics of the surgical situation that are identified by most patients as stressful. However, it seems likely that different types of surgical procedures produce different types of stress. A useful way of distinguishing between the various procedures is by considering the function of the procedure (diagnostic, treatment, or both) and the time line and nature of stress associated with the procedure. A further distinction can be drawn between procedural stress (i.e., the stress associated with the negative aspects of the actual procedure itself) and outcome stress (i.e., longer-term fears and concerns related to the results of the treatment or procedure). To illustrate the latter point, some operations may have more positive characteristics in terms of their expected outcome than others, for example, restoration (hip replacement) versus the removal of physical function (leg amputation).

Responses of Surgical Patients at Emotional, Cognitive, and Physiological Levels

Emotional Responses

Most studies investigating emotional responses to surgery have shown elevated levels of anxiety both before and after surgery. In some groups of patients, postoperative anxiety levels may be even higher than those measured preoperatively. Previous results indicating that anxiety is high before but not after surgery have been dismissed as being methodically flawed. Although anxiety has been, for obvious reasons, the most frequently assessed emotional response, it has been shown that surgical patients may also experience high levels of nervousness, depression, anger, and boredom.

In summary, the literature suggests that emotional responses to the surgical situation may vary as a function of the type of operation and, therefore, outcome concerns. Study results also point to the possibility of effective interventions not only before but also after surgery.

Cognitive Responses

Several studies have shown that patients' main worries are more related to the outcome of the operation than to the operation itself. Patients also worry significantly about normal everyday matters such as family and home, perhaps exacerbated by hospitalization and the impending surgery. As in the treatment of depression, the cognitive elements of surgical patients' stress response may be critical in indicating the most promising intervention approach to alleviate distress.

Physiological Responses

The most thorough research on physiological responses to surgery has been carried out on indices of sympathetic-adrenomedullary activity, both because they are readily measurable in surgical patients and because this highly responsive physiological axis has been widely used in well-controlled laboratory studies of psychological stress. The most consistent finding is that of a reduction in palmar sweat gland activity prior to surgery followed by recovery to normal levels postoperatively. This pattern follows closely the pattern of subjective distress and pain, indicating that sweat gland activity is reduced by the stress of surgery. Some authors, however, have argued that the palmar sweating pattern of surgical patients is related to the effort rather than to the distress aspect of stress, that is, that changes in palmar sweating are due to changes in activity levels.

A more recent line of research has investigated neuroendocrine and immune changes in response to surgery and wound healing. Postoperative elevations in plasma levels of adrenaline, cortisol, and β-endorphins reflect sympathetic nervous system and hypothalamic-pituitary-adrenal axis activation. Evidence for immune suppression during surgery comes from studies showing the suppression of natural killer cell activity, proliferation of lymphocyte responses to mitogens, and changes in lymphocyte populations.

In order to interpret physiological responses to surgery as psychophysiological phenomena, it is necessary to disentangle the effects that are due to the experience of stress from those that are caused by the surgical trauma and other medical interventions (e.g., anesthesia). Apart from anesthesia and surgery itself, physiological changes may be caused by alterations in diet and ambulation and by medications, especially starvation before surgery and premedication or sedation on the night prior to surgery. Although these problems have at least been partly addressed in studies of the sympathetic-adrenomedullary axis, studies of other pathways are still at a very early stage and cannot support any firm conclusions.

Surgical Stress and Postoperative Recovery

Despite a strong effect of physical factors such as extent of tissue damage caused by the surgical trauma on postoperative recovery, there is variability across patients who have undergone the same procedure. It is unclear what accounts for this variability, and this has led to the suggestion that psychological factors such as anxiety and depression may play some part in determining the duration or quality of postoperative recovery.

There is considerable theoretical and empirical support for the relevance of psychosocial factors to postoperative recovery. These include demographic variables, such as age, gender, and socioeconomic status, and also personality variables, such as anxiety, neuroticism, coping style and social support. Most of the available literature today suggests a linear association between preoperative anxiety and success of recovery, in that the more anxious the patient is before the operation, the less favorable the postoperative recovery. There is even evidence that preoperative anxiety is related to intraoperative adjustment.

There is increasing evidence from psychoneuroimmunological work that stress and pain have a negative effect on endocrine and immune functions and

may thereby lead to delays in wound healing and other clinical complications. Although wound healing is only one postoperative outcome among a whole range, it is certainly a central parameter for short-term outcome in recovery from surgery. Kiecolt-Glaser et al. offer a biobehavioral model suggesting a number of routes through which psychological and behavioral responses can influence surgery and postoperative recovery.

Psychological Preparation for Surgery

There is now substantial agreement that psychological preparation for surgery is beneficial for patients. It is important, however, to establish which benefits can be achieved by psychological preparation and if all forms of preparation are equally effective. Results from a number of meta-analyses suggest that significant benefits can be obtained on all the major outcome variables (i.e., negative affect, pain, pain medication, length of stay, behavioral and clinical indices of recovery, physiological indices, and satisfaction). The methods of psychological preparation employed include the provision of information about sensations and procedures, behavioral instructions, cognitive interventions, relaxation, and hypnosis. Procedural information and behavioral instructions show the most ubiquitous effects, both being clearly effective in improving all postoperative outcomes.

Further Reading

Johnston, M. and Vögele, C. (1993). Benefits of psychological preparation for surgery: a meta-analysis. *Annals of Behavioral Medicine* **15**, 245–256.

Kiecolt-Glaser, J. K., Page, G. G., Marucha, P. T., et al. (1998). Psychological influences on surgical recovery: perspectives from psychoneuroimmunology. *American Psychologist* **53**, 1209–1218.

Vögele, C. (2004). Hospitalisation and stressful medical procedures. In: Kaptein, A. & Weinman, J. (eds.) *Health psychology*, pp. 288–304. Oxford: BPS Blackwell Publishers.

Survivor Guilt

P Valent
Melbourne, Victoria, Australia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Valent, volume 3, pp 555–557, © 2000, Elsevier Inc.

History of the Concept of Survivor Guilt
Moral Features of Survivor Guilt
Sense and Purpose of Survivor Guilt
Range of Manifestations of Survivor Guilt
Treatment of Survivor Guilt

Glossary

Guilt	An internal emotional judgment (an aspect of conscience) of having been bad or having done wrong to another.
Survivor	A person who has lived through a trauma.
Trauma	An experience in which a person's life has been grossly threatened and out of which a variety of biological, psychological and social wounds and scars results.

History of the Concept of Survivor Guilt

Survivor guilt as a clinical entity was emphasized for the first time in Holocaust literature in the 1960s. Psychiatrists such as Krystal and Niederland were struck with the ubiquity of feelings of intense, unabating guilt for having survived among their Holocaust survivor patients. Survivor guilt came to be quickly recognized in other traumatic situations, including the Hiroshima atom bomb, combat, disasters, and civilian bereavement. It was noted after 9/11; indeed, it can occur in all traumatic situations, whether natural or man made.

A special form of survivor guilt may be said to occur in rescuers and helpers who blame themselves for not having saved those for whom they viewed themselves as responsible. More subtle forms can occur in clinicians who feel guilt for their own well-being in the face of their patients' suffering.

Moral Features of Survivor Guilt

The most striking feature of survivor guilt is its moral judgment of self-blame. It is directed to the survivor's

undeserved survival in contrast to others' deaths. An individual suffering from survivor guilt often states that the dead should be alive instead of him- or herself. This feeling may be enhanced when the individual was saved by another who died.

A common subtype of survivor guilt relates to not having saved others. The implication is that the survivor could have and should have saved others, but due to selfishness did not, instead saving him- or herself, thereby causing suffering and death to others. This feeling is enhanced if the person's usual role, such as spouse, parent, or rescuer, was to protect those who died, or if the person did indeed put his or her survival ahead of others.

However, a striking feature of survivor guilt is the contrast between the torment of self-blame of survivors and their innocence by any objective criteria. For instance, Holocaust survivors often blamed themselves for not following or saving their loved ones who were murdered, when actually they were separated at gunpoint and through physical force. Children who survived by hiding blamed themselves for the deaths of their parents. The Holocaust taught the paradox that victims are prone to survivor guilt while perpetrators rationalize theirs.

Unjustified survivor guilt occurs in all traumatic situations. Medical staff may blame themselves for not saving unsavable patients. A child may grow up with unwarranted guilt for separation from parents or parental death.

Sense and Purpose of Survivor Guilt

Observations of acute phases of traumatic situations indicate that survivor guilt is an unpleasant, instinctive social cue (that may be reinforced by others' anger and blame) that can be appeased only through helping others. For instance, in a disaster those whose houses survive often appease their very intense guilt by sheltering those whose houses are destroyed. Survivor guilt thus helps to preserve as many people (or genes) in the community as possible.

Later survivor guilt may be less useful practically, though it may prime adaptive reparative action for subsequent traumatic situations. It may also preserve the survivor's retrospective hope that 'if onlys' may yet be able to be executed. Further, Danieli suggests that the guilt is a buffer against moral chaos and existential helplessness, as it preserves some assumption of moral order in the universe. Last, the guilt perpetuates close emotional links to the dead.

Range of Manifestations of Survivor Guilt

If survivor guilt is appeased by effective rescue of others, it may be seen as a temporary stress. If it fails and become part of a trauma, it is lived and relived as part of the wounds and scars of the trauma. Then the following biological, psychological, and social manifestations may occur singly or in a variety of combinations.

Physiological Responses

Thus far, the physiological accompaniments of survivor guilt have not been delineated.

Psychological Responses

Emotion Survivor guilt is one of the most painful human emotions. It may be felt as both mental anguish and a physical heart-wrenching pain in the chest directed toward the dead, but turned angrily back on the survivor him- or herself.

Cognitions Poignant traumatic events may return in feelings, thoughts, images, dreams, and flashbacks, along with ruminations about what the survivor should have done but did not do. Cognitive schemas and meanings develop, such as that the survivor is an irresponsible person, even a destroyer of life.

Defenses Intense pain of survivor guilt such as having caused the death of beloved peoples, may be central in survivor defense formations. Psychic numbing, dissociation, and repression may fragment coherent awareness. Identification may lead to physical symptoms similar to the ones the dead had suffered. Displacement may lead to blaming others. Sublimation may lead to devotion to saving others. Some survivors may take on rescuer or helper professions.

Social Responses

In acute situations, people may vacillate between appeasing survivor guilt by helping others and other essential survival options. After the acute phase of the trauma, people may wring their hands as they anguish about the choices they made. They may blame themselves and withdraw in order to avoid the imagined blame of the question "Why did you survive?" Survivors may play down their survivorship, and may especially avoid enjoyment and vitality, as they can exacerbate their feelings of guilt. Sublimatory activity may be channeled through choosing helping or rescuing work and professions.

Specific social settings may lead to characteristic responses. For instance, bereaved parents often withdraw to extremes, and emotional numbing admixed with intense blame and self-blame often result in divorce.

Associated Clinical Features

Survivor guilt may be associated with other judgments, such as shame and injustice. The sense of

injustice may come from having been unfair to others, but also from having been put in unfair situations.

Survivor guilt may interfere with grieving and is often at the kernel of unresolved grief and depression. Hence it may need to be resolved before the grieving process can proceed.

Treatment of Survivor Guilt

Prevention of survivor guilt is part of the rationale for strict, following of protocols, early disaster intervention, debriefing, and grief therapy. Its resolution is a common goal in later psychotherapies.

The first principle of treatment of survivor guilt is its recognition. Next, thorough investigation of the facts of the circumstances reveals how survivor guilt is objectively untenable. Its evolutionary sense and hence its presence is explained, and alternative, hopeful cognitive views are explored. With the thawing defenses of the survivor, the emotional pain of survivor guilt is identified, processed, and assimilated in the context of an empathic, therapeutic relationship. The survivor's sense of morality is retrieved with the realization that past circumstances, not the survivor, were irrational and bad. Seeing him- or herself as a victim, not as a perpetrator, allows the survivor to grieve losses and to achieve new hopeful meanings for him- or herself. The survivor can then reclaim a purposeful life.

See Also the Following Articles

Depression and Manic-Depressive Illness; Holocaust, Stress Effects of; 9/11, Religion and Stress.

Further Reading

Danieli, Y. (1988). Treating survivors and children of the Nazi Holocaust. In: Ochberg, F. M. (ed.) *Post-traumatic therapy and victims of violence*. New York: Brunner/Mazel.

Krystal, H. and Niederland, W. C. (eds.) (1971). *Psychic traumatization: aftereffects in individuals and communities*. Boston, MA: Little, Brown and Company.

Raphael, B. (1986). *When disaster strikes*. London: Hutchinson.

Valent, P. (1998). *From survival to fulfillment: a framework for the life-trauma dialectic*. Philadelphia, PA: Brunner/Mazel.

Sympathetic Nervous System

D S Goldstein
National Institute of Neurological Disorders and Stroke, Bethesda, MD, USA

Published by Elsevier Inc.

Historical and Conceptual Overview
Role of the Sympathetic Nervous System in Homeostasis
Norepinephrine: The Sympathetic Neurotransmitter

Glossary

Sympathetic nervous system	A major component of the autonomic nervous system. The main sympathetic neurotransmitter for cardiovascular regulation is norepinephrine. The main sympathetic neurotransmitter for sweating is acetylcholine.
Adrenomedullary hormonal system	The hormonal component of the autonomic nervous system.
Norepinephrine	The main sympathetic neurotransmitter for cardiovascular regulation.
Epinephrine	The main hormone of the adrenomedullary hormonal system.
Norepinephrine	The sympathetic neurotransmitter.

Historical and Conceptual Overview

The term sympathetic nervous system originates in the second-century teaching of Galen that the peripheral nerves, conduits for distributing the animal spirit in the body, enable concerted, coordinated (i.e., sympathetic) functioning of body organs. The notion of the sympathetic nervous system, therefore, antedated by approximately 14 centuries Harvey's demonstration of the circulation of the blood.

The sympathetic nervous system is a major component of the autonomic nervous system. In the late nineteenth century, Langley introduced the phrase autonomic nervous system to refer to nerves emanating from cell bodies in ganglia alongside the spinal column and in the gastrointestinal-tract walls. Langley defined three components: sympathetic, with preganglionic

cells in the thoracolumbar spinal cord; parasympathetic (a word he coined), with preganglionic cells in the brain stem or sacral spinal cord; and enteric, with preganglionic cells near or in the target organs. The nerves of these systems were thought to differ from those of the somatic nervous system, not only in terms of anatomy but also in terms of function. Somatic nerves mediate conscious, observable changes in skeletal muscle contraction and therefore movement, whereas autonomic nerves mediate unconscious, largely unobservable changes in smooth muscle contraction independently (autonomously) of the central nervous system.

The early-twentieth century American physiologist Walter B. Cannon introduced the word homeostasis. Cannon added a hormonal component to the autonomic nervous system when he theorized that the sympathetic nervous system and the adrenal gland worked together as a unit to maintain homeostasis in emergencies. In the 1930s, he even formally proposed that the sympathetic nervous system used the same chemical messenger – adrenaline – as did the adrenal gland (in 1946, von Euler correctly identified the sympathetic neurotransmitter in mammals as norepinephrine). Accumulating evidence supports the independent regulation of the sympathetic nervous and adrenomedullary hormonal systems, refuting the concept of a unitary sympathoadrenal system. Nevertheless, Cannon's views about the unitary function of the neural and hormonal components of the sympathoadrenal system prevail in stress research.

Sweating, whether to control body temperature (thermoregulatory), evoked by eating spicy food (gustatory), or accompanying fear or anxiety (emotional), is mediated by the sympathetic cholinergic system, in which acetylcholine rather than norepinephrine is the chemical messenger. The autonomic nervous system, therefore, can be viewed as having five components: sympathetic noradrenergic, sympathetic cholinergic, parasympathetic cholinergic, adrenomedullary, and enteric.

Role of the Sympathetic Nervous System in Homeostasis

Adjustments in sympathetic nervous activity accompany virtually every human action and reaction and constitute the main means for the acute regulation of cardiovascular function. According to Goldstein's 1995 theory, comparator homeostats use the sympathetic nervous system as an effector for different stress responses. The sympathetic nervous and adrenomedullary hormonal effectors interact in complex ways with many others, including the hypothalamic–pituitary–adrenocortical system, the endogenous opiate system, vasopressin, the parasympathetic nervous system, and the renin–angiotensin–aldosterone system, in determining responses to stressors.

The sympathetic nervous system consists of nerve networks. Because the axons come from cells in ganglia rather than directly from the spinal cord or brain stem, sympathetic nerves supplying body organs are mainly postganglionic; however, most sympathetic nerves supplying adrenomedullary chromaffin cells are preganglionic, releasing acetylcholine and thereby stimulating membrane nicotinic receptors.

Stresses associated with sympathetic nervous system activation typically include a component of active movement. Patterned sympathetic nervous system activation during stress produces adaptive shifts in the distribution of blood volume or in glandular secretion (**Figure 1a**). When these changes suffice to maintain homeostasis, the organism does not experience them consciously, but when the brain senses that these responses are not or will not mitigate the effects of the stressor, the situation reaches consciousness and adrenomedullary hormonal system activation ensues (**Figure 1b**). Examples of situations associated with prominent increases in sympathoneural outflows include orthostasis, exercise, the postprandial state, cold exposure, and the performance of nondistressing locomotor tasks.

In response to perceived global metabolic threats from external or internal stimuli, increased neural outflow to the adrenal medulla elicits catecholamine secretion into the adrenal venous drainage. In humans, the predominant catecholamine in the adrenal venous drainage is epinephrine (synonymous with adrenaline). Epinephrine rapidly reaches all cells of the body, with the exception of most of the brain, producing a wide variety of hormonal effects at low blood concentrations. We can comprehend the many effects of epinephrine in terms of the body's countering of acute threats to survival that mammals have perennially faced, such as sudden lack of metabolic fuels, trauma with hemorrhage, intravascular volume depletion and hypotension, and fight-or-flight confrontations. Distress accompanies all these situations, the experience undoubtedly fostering the long-term survival of the individual and the species by motivating avoidance learning, enhancing the registration of the event in long-term memory, and producing signs universally understood among other members of the species.

Stresses eliciting adrenomedullary hormonal system activation typically also elicit hypothalamic-pituitary-adrenocortical activation, as indicated by high circulating levels of corticotropin or glucocorticoids; increases in release of endogenous opioids, as indicated by high plasma levels of β-endorphin; and

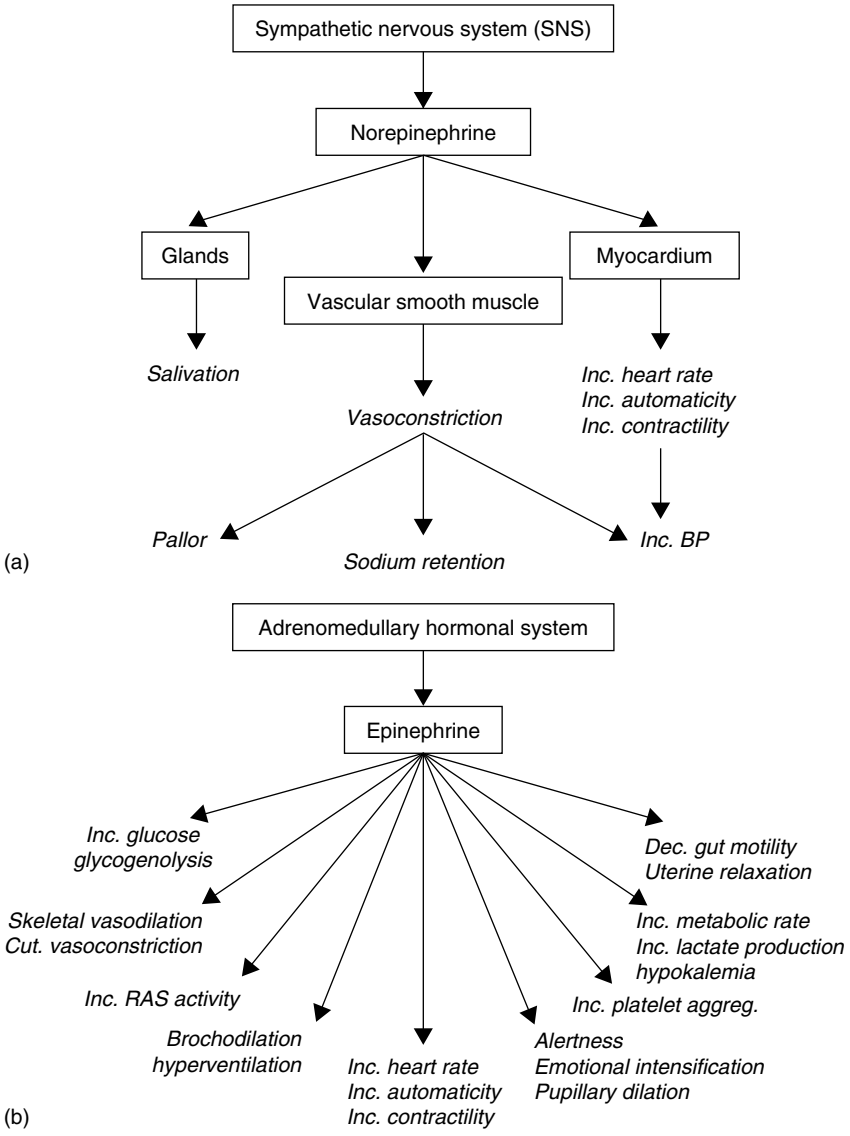


Figure 1 Overview of effects of the activation of (a) the sympathetic nervous system, where norepinephrine acts as the main neurotransmitter, and (b) the adrenomedullary hormonal system, where epinephrine acts as the main hormone. Because the sympathetic nervous system is a neural network, sympathetic nerves to different regions can be activated differentially during stress, producing redistributions of blood flow to body organs. Norepinephrine released from sympathetic nerves generally contracts smooth muscle cells, eliciting glandular secretion, vasoconstriction, and myocardial contraction. Renal vasoconstriction produces an anti-natriuretic effect. Cutaneous vasoconstriction produces pallor. The net effect of systemic vasoconstriction is to increase the total peripheral resistance to blood flow in the body. This, combined with myocardial stimulation, increases blood pressure. In humans, the main hormone released by the adrenal medulla is epinephrine (adrenaline). Epinephrine affects the function of most organs and produces generalized effects such as increased metabolic rate, increased blood glucose levels, hypokalemia, increased lactate production, increased alertness, increased platelet aggregability, and intensification of emotions. Aggreg., aggregation; BP, blood pressure; Cut., cutaneous; Dec., decrease; Inc., increase; RAS, renin–angiotensin–aldosterone system.

vasopressin secretion even when sympathetic nervous system activity declines.

Norepinephrine: The Sympathetic Neurotransmitter

The chemical transmitter at most cardiovascular sympathetic nerve endings is norepinephrine. The binding of norepinephrine to adrenoceptors on cardiovascular smooth muscle cells contracts the cells. The body's myriad arterioles largely determine total resistance to blood flow and therefore contribute in an important way to blood pressure. Sympathetic nerves enmesh blood vessels in latticelike networks in the adventitial outer surface that extend inward to the adventitial-medial border, with the concentration of sympathetic

nerves increasing as arterial caliber decreases, so that small arteries and arterioles, the smallest nutrient vessels possessing smooth muscle cells, possess the most intense innervation. The unique architectural association between sympathetic nerves and the vessels that determine peripheral resistance has enticed cardiovascular researchers, particularly in the area of autonomic regulation, for many years. Sympathetic vascular innervation varies widely among vascular beds, with dense innervation of resistance vessels in the gut, kidney, skeletal muscle, and skin. Sympathetic stimulation in these beds produces profound vasoconstriction, whereas stimulation in the coronary, cerebral, and bronchial beds elicits weaker constrictor responses, consistent teleologically with the goal of preserving blood flow to vital organs during stress. Cardiac sympathetic stimulation increases the force and rate of the heartbeat and also augments intracardiac electrical conduction and automaticity.

Norepinephrine released from sympathetic nerve terminals acts mainly locally. Only a small proportion of released norepinephrine reaches the bloodstream. The indirect and distant relationship between plasma norepinephrine levels and sympathetic nerve activity can complicate interpreting plasma norepinephrine levels in response to stressors, pathophysiological situations, and drugs.

Norepinephrine Synthesis

The enzymatic steps in norepinephrine synthesis have been characterized in more detail than those for any other neurotransmitter (Figure 2). Catecholamine

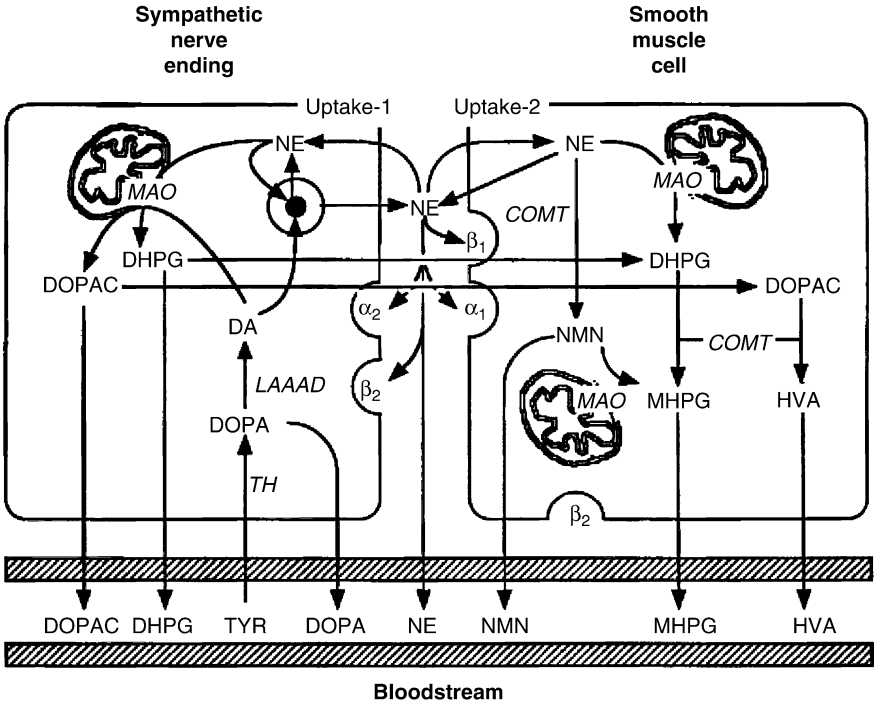


Figure 2 Overview of steps in norepinephrine synthesis, release, reuptake, and metabolism. Hydroxylation of the amino acid tyrosine is the enzymatic rate-limiting step in catecholamine biosynthesis. Tyrosine hydroxylase catalyzes the conversion of tyrosine to DOPA. L-aromatic-amino acid decarboxylase (LAAAD) catalyzes the conversion of DOPA to dopamine. Dopamine-β-hydroxylase is localized to large dense-core vesicles in noradrenergic cells. After uptake of axoplasmic dopamine into the vesicle, dopamine-β-hydroxylase catalyzes the conversion of dopamine to norepinephrine. According to the exocytosis theory, increased transmembrane entry of ionized calcium fosters fusion of the vesicles with the axonal membrane. Endogenously released norepinephrine is taken back up into sympathetic nerve terminals by Uptake-1 and taken up into nonneuronal cells by Uptake-2. Norepinephrine in the axoplasm can be translocated back into the vesicles or converted to dihydroxyphenylglycol by monoamine oxidase (MAO) in the outer mitochondrial membrane. Norepinephrine in nonneuronal cells is converted to normetanephrine by catechol-O-methyltransferase. Dihydroxyphenylglycol can diffuse into the bloodstream or can be taken up by nonneuronal cells and converted to methoxyhydroxyphenylglycol, which is a major end product of norepinephrine metabolism. Dopamine in the axoplasm can be translocated into vesicles and converted to norepinephrine or can be converted to dihydroxyphenylacetic acid by MAO. Dihydroxyphenylacetic acid can diffuse into the bloodstream or can be converted extraneuronally to homovanillic acid, the main end product of dopamine metabolism. The processes of catecholamine uptake and metabolism are generally independent of actions by adrenoceptors on the nerve terminals or nonneuronal cells. COMT, catechol-O-methyltransferase; DA, dopamine; DHPG, dihydroxyphenylglycol; DOPA, 3,4-dihydroxyphenylalanine; DOPAC, dihydroxyphenylacetic acid; HVA, homovanillic acid; LAAAD, L-aromatic-amino acid decarboxylase; MAO, monoamine oxidase; MHPG, methoxyhydroxyphenylglycol; NE, norepinephrine; NMN, normetanephrine; TH, tyrosine hydroxylase; TYR, tyrosine.

biosynthesis begins with the uptake of the amino acid tyrosine into the cytoplasm of sympathetic neurons and adrenomedullary cells. Tyrosine hydroxylase catalyzes the conversion of tyrosine to 3,4-dihydroxyphenylalanine (DOPA). This is the enzymatic rate-limiting step in catecholamine synthesis. The enzyme is almost saturated under normal conditions, so alterations in dietary tyrosine intake normally do not affect the rate of catecholamine biosynthesis under baseline conditions. Tetrahydrobiopterin, ionic iron, and molecular oxygen are required for tyrosine hydroxylase activity. Exposure to stressors that increase sympathoadrenal outflows augments the synthesis and concentration of tyrosine hydroxylase by multiple and complex mechanisms. Short-term mechanisms include feedback inhibition and phosphorylation of the enzyme, the latter depending on membrane depolarization, contractile elements, and receptors. Long-term mechanisms include changes in tyrosine hydroxylase synthesis, probably determined by transsynaptic induction and retrograde transport.

L-aromatic-amino-acid decarboxylase (LAAAD, also called DOPA decarboxylase) in the neuronal cytoplasm catalyzes the rapid conversion of DOPA to dopamine. Many types of tissues contain this enzyme. The activity of the enzyme depends on pyridoxal phosphate, which is vitamin B-6. Although LAAAD metabolizes most of the DOPA formed in catecholamine-synthesizing tissues, some endogenously produced DOPA enters the circulation. This provides the basis for using plasma DOPA levels to examine catecholamine synthesis. By inhibiting the conversion of DOPA to dopamine in the periphery, LAAAD inhibitors such as carbidopa enhance the efficacy of L-DOPA treatment of Parkinson's disease.

Dopamine-β-hydroxylase catalyzes the conversion of dopamine to norepinephrine in tissues that synthesize catecholamines. The enzyme is confined to the vesicles in catecholamine-producing cells. Dopamine-β-hydroxylase contains, and its activity depends on, copper. Because of this dependence, children with Menkes disease, a rare, inherited X-linked recessive disorder of copper metabolism, have neurochemical evidence of decreased conversion of dopamine to norepinephrine and compensatorily increased tyrosine hydroxylation. Dopamine-β-hydroxylase activity also requires ascorbic acid (vitamin C), which provides electrons for the hydroxylation.

Norepinephrine Storage

Vesicles in sympathetic nerves actively remove and trap axoplasmic amines. The vesicular monoamine transporter resembles the neuronal membrane monoamine transporters structurally, but it is relatively nonspecific for catecholamines and serotonin. Magnesium and ATP accelerate vesicular uptake, and reserpine effectively and irreversibly blocks it. Vesicles generated near the Golgi apparatus of the cell bodies travel by axonal transport in the axons to the nerve terminals. Noradrenergic vesicles may also form by endocytosis within the axons. Varicosities in sympathetic nerves contain two types of cytoplasmic vesicles: small dense-core (diameter 40–60 nm) and large dense-core (diameter 80–120 nm). Cores of both types of vesicle contain abundant ATP, the norepinephrine : ATP ratio averaging approximately 1 : 4. The vesicles also contain at least three types of polypeptides: chromogranin A, enkephalins, and neuropeptide Y. The extracellular fluid levels of each of these compounds have been considered as putative indices of exocytosis.

Norepinephrine Release

According to the exocytotic theory, acetylcholine depolarizes the terminal membranes by increasing membrane permeability to sodium ion. The increased intracellular sodium levels directly or indirectly enhance the transmembrane influx of calcium ion via voltage-gated channels. The increased cytoplasmic calcium concentration evokes a cascade of as yet incompletely defined biomechanical events, resulting in the fusion of the vesicular and axoplasmic membranes. The interior of the vesicle exchanges briefly with the extracellular compartment, and the soluble contents of the vesicles diffuse into the extracellular space. As predicted from this model, manipulations other than the application of acetylcholine that depolarize the cell, such as electrical stimulation or increased potassium ion concentrations in the extracellular fluid, also activate the voltage-gated calcium channels and trigger exocytosis.

During sympathetic neuronal or adrenomedullary chromaffin cell activation, the simultaneous stoichiometric release of soluble vesicular contents – ATP, enkephalins, chromogranins, and dopamine-β-hydroxylase – without a similar release of cytoplasmic macromolecules has provided biochemical support for the exocytosis theory of noradrenergic neurotransmission. Electron micrographs occasionally have shown an omega sign, with an apparent gap in the cell membrane at the site of fusion of vesicle with the axoplasmic membrane. Ultrastructural evidence has suggested the release of the contents of large as well as small dense-core vesicles by exocytosis.

Sympathetic nerve endings can also release norepinephrine by calcium-independent, nonexocytotic mechanisms. One such mechanism probably is reverse transport through the neuronal uptake carrier. The indirectly acting sympathomimetic amine, tyramine,

releases norepinephrine nonexocytotically because tyramine releases norepinephrine independently of calcium and does not release dopamine-β-hydroxylase. As predicted from the ionic requirements for neuronal reuptake, increases in intracellular sodium concentrations, such as produced by ouabain, enhance carrier-mediated efflux of norepinephrine. Myocardial ischemic hypoxia evokes the calcium-independent release of norepinephrine. The hydrophilic nature of catecholamines and their ionization at physiological pH probably prevent norepinephrine efflux by simple diffusion.

Pharmacological stimulation of receptors on noradrenergic terminals affects the amount of norepinephrine released during cellular activation. Substantial evidence supports inhibitory presynaptic modulation by norepinephrine itself via α2-adrenoceptors on sympathetic nerves. This modulatory action appears to vary with the vascular bed under study, being prominent in skeletal muscle beds such as the forearm, relatively weak in the kidneys, and virtually absent in the adrenals. In addition to local feedback control of norepinephrine release, reflexive long-distance feedback pathways via arterial and cardiac baroreceptors elicit reflexive changes in sympathoneural impulse activity. Alterations in receptor numbers or of intracellular biomechanical events after receptor activation also affect responses to agonists. These factors therefore regulate norepinephrine release by trans-synaptic local and reflexive long-distance mechanisms.

Norepinephrine Disposition

Unlike acetylcholine, which is inactivated mainly by extracellular enzymes, norepinephrine is inactivated mainly by uptake into cells (**Figure 2**), with subsequent intracellular metabolism or storage. Reuptake into nerve terminals (Uptake-1) is the predominant means of terminating the actions of released norepinephrine. Dopamine is a better substrate than norepinephrine is for uptake via the cell-membrane norepinephrine transporter.

Neuronal uptake absolutely requires intracellular potassium ion and extracellular sodium ion and functions most efficiently when chloride ion accompanies the sodium. Uptake-1 requires energy and is carrier mediated. Transport does not directly require ATP; however, maintaining ionic gradients across cell membranes depends on ATP, and the carrier uses the energy expended in maintaining the transmembrane sodium gradient to cotransport amines with sodium. Many drugs or *in vitro* conditions inhibit Uptake-1, including cocaine and tricyclic antidepressants.

Norepinephrine taken up into the axoplasm by the Uptake-1 transporter is subject to two fates: translocation into storage vesicles and deamination by monoamine oxidase (MAO). The combination of enzymatic breakdown and vesicular uptake constitute an intraneuronal sink, keeping the cytoplasmic concentrations of norepinephrine very low. Reserpine effectively blocks the vesicular transport of amines from the axoplasm into the vesicles. This not only shuts down the conversion of dopamine to norepinephrine but also prevents the conservative recycling of norepinephrine. Reserpine, therefore, rapidly depletes norepinephrine stores.

Neural and nonneural tissues contain MAO, which catalyzes the oxidative deamination of dopamine to form dihydroxyphenylacetic acid and norepinephrine to form dihydroxyphenylglycol. Because of the efficient uptake and reuptake of catecholamines into the axoplasm of catecholaminergic neurons and because of the rapid exchange of amines between the vesicles and axoplasm, the neuronal pool of MAO, located in the outer mitochondrial membrane, figures prominently in the overall functioning of catecholaminergic systems. The isozyme MAO-A predominates in neural tissue, whereas both MAO-A and MAO-B exist in nonneuronal tissue. MAO inhibitors are effective antidepressants. A phenomenon known as the cheese effect limits their clinical use. In patients taking MAO inhibitors, the administration of sympathomimetic amines such as those found in many nonprescription decongestants or the ingestion of foods such as aged cheese, wine, and meat, which contain tyramine, can produce paroxysmal hypertension. Tyramine and other sympathomimetic amines increase the loss of norepinephrine from sympathetic vesicles, and the blockade of MAO augments the bioavailability of ingested tyramine. When given alone, MAO inhibitors usually decrease blood pressure and can produce orthostatic hypotension by unknown mechanisms.

Nonneuronal cells remove norepinephrine actively by a process called Uptake-2, characterized by the ability to transport isoproterenol; susceptibility to blockade by O-methylated catecholamines, corticosteroids, and β-haloalkylamines; and an absence of susceptibility to blockade by Uptake-1 blockers. In contrast with Uptake-1, Uptake-2 functions independently of extracellular sodium. The Uptake-2 carrier has little if any stereoselectivity and has low affinity and specificity for catecholamines. For instance, extraneuronal cells remove imidazolines such as clonidine by Uptake-2. Whereas reverse transport via the Uptake-1 carrier requires special experimental conditions, reverse transport via the Uptake-2 carrier is

readily demonstrated. Thus, during the infusion of a catecholamine at a high rate, the catecholamine can accumulate in extraneuronal cells, with reentry of the catecholamine into the extracellular fluid via the Uptake-2 carrier after the infusion ends.

Catechol-O-methyltransferase in nonneuronal cells and adrenomedullary chromaffin cells catalyzes the O-methylation of dihydroxyphenylglycol to form methoxyhydroxyphenylglycol and of dihydroxyphenylacetic acid to form homovanillic acid. Plasma levels of dihydroxyphenylglycol and dihydroxyphenylacetic acid probably mainly reflect neuronal metabolism of norepinephrine and dopamine. Catechol-O-methyltransferase also catalyzes the conversion of norepinephrine to normetanephrine and epinephrine to metanephrine. Uptake-2 and catechol-O-methyltransferase probably act in series to remove and degrade circulating catecholamines. The methyl group donor for the reaction is S-adenosyl methionine. Immunohistochemical studies have indicated extraneuronal localization of catechol-O-methyltransferase in most tissues, such as the liver and kidney. Because of the extraneuronal localization of catechol-O-methyltransferase, recently introduced assay methods for metanephrines have enabled refined assessments of extraneuronal metabolism of catecholamines.

Most of the metabolism of catecholamines takes place within the same cells where the amines are synthesized. This mainly occurs secondary to leakage of catecholamines from vesicular stores into the cytoplasm. The stores exist in a highly dynamic equilibrium, with passive outward leakage counterbalanced by inward active transport controlled by vesicular monoamine transporters.

In catecholaminergic neurons, the presence of MAO-A leads to the formation of reactive catecholaldehydes. The contribution of these toxic aldehydes to neurodegenerative processes depends on the dynamics of vesicular-axoplasmic monoamine exchange and enzyme-catalyzed conversion to inert deaminated acids or alcohols. In sympathetic nerves, the aldehyde produced from norepinephrine is converted to dihydroxyphenylglycol, not dihydroxymandelic acid, as textbooks often indicate. Subsequent extraneuronal O-methylation consequently leads to the production of methoxyhydroxyphenylglycol, not vanillylmandelic acid. Vanillylmandelic acid is instead formed in the liver by the oxidation of methoxyhydroxyphenylglycol catalyzed by alcohol dehydrogenase. Compared to intraneuronal deamination, extraneuronal O-methylation of norepinephrine and epinephrine to metanephrines represent minor pathways of metabolism. The single largest source of metanephrines is the adrenal medulla. Vanillylmandelic acid and methoxyhydroxyphenylglycol, the products of the combined O-methylation and deamination of norepinephrine, are the two main end products of norepinephrine metabolism, vanillylmandelic acid (as just noted) formed mainly in the liver. Adrenomedullary chromaffin cells express catechol-O-methyltransferase, as do pheochromocytomas, which are clinically important chromaffin cell tumors. Pheochromocytoma cells produce large amounts of metanephrines from catecholamines leaking from stores. These metabolites are thus particularly useful for detecting these tumors. The large contribution of intraneuronal deamination to catecholamine turnover, and the dependence of this on the vesicular axoplasmic monoamine exchange process, help explain how synthesis, release, metabolism, turnover, and stores of catecholamines are regulated in a coordinated fashion.

See Also the Following Article

Autonomic Nervous System.

Further Reading

Cannon, W. B. (1939). *The wisdom of the body*. New York: W. W. Norton.

Eisenhofer, G., Kopin, I. J. and Goldstein, D. S. (2004). Catecholamine metabolism: a contemporary view with implications for physiology and medicine. *Pharmacological Reviews* **56**, 331–349.

Goldstein, D. S. (1995). *Stress, catecholamines, and cardiovascular disease*. New York: Oxford University Press.

Goldstein, D. S. (2001). *The autonomic nervous system in health and disease*. New York: Marcel Dekker.

Goldstein, D. S., Eisenhofer, G. and Kopin, I. J. (2003). Sources and significance of plasma levels of catechols and their metabolites in humans. *Journal of Pharmacology and Experimental Therapeutics* **305**, 800–811.

Pacak, K., Palkovits, M., Yadid, G., et al. (1998). Heterogeneous neuroendocrine responses to various stressors: a test of Selye's doctrine of non-specificity. *American Journal of Physiology* **275**, R1247–R1255.

Selye, H. (1950). *The physiology and pathology of exposure to stress: a treatise based on the concepts of the general-adaptation syndrome and the diseases of adaptation*. Montreal: Acta, Inc.

Syndrome X *See:* Metabolic Syndrome; Metabolic Syndrome and Stress.

Synthetic Glucocorticoids

A M Karssen
Stanford University School of Medicine, Stanford, CA, USA
E R de Kloet
Leiden University Medical Center, Leiden, The Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A M Karssen and E R de Kloet, volume 3, pp 566–570, © 2000, Elsevier Inc.

Introduction
Chemical Aspects
Pharmacokinetics
Actions of Synthetic Glucocorticoids
Clinical Aspects

Glossary

Blood–brain barrier	Dynamic barrier of endothelial cells between blood and brain, protecting the brain from potentially neurotoxic compounds and maintaining a constant internal environment in the brain.
Corticotropin releasing hormone (CRH)	A 41 amino acid residue peptide that is synthesized in hypothalamic nerve cells and mediates the neural control of ACTH release.
Glucocorticoids	Steroid hormones that affect carbohydrate and protein metabolism, regulate the activity of the hypothalamic-pituitary-adrenal (HPA) axis, and have anti-inflammatory and immunosuppressive effects. Endogenously, they are synthesized in the adrenal cortex, released on stimulation by adrenocorticotropic hormone (ACTH).
Mineralocorticoids	Steroid hormones that affect water and electrolyte balance. Naturally occurring glucocorticoids often have also substantial mineralocorticoid action.

Introduction

Synthetic glucocorticoids are widely used in basic, preclinical, and clinical research. The importance of these compounds in stress research is based on their suppressive effects on the activity of the hypothalamic-pituitary-adrenal axis (HPA axis). Their main clinical applications, however, are based on their effective anti-inflammatory and immunosuppressive actions. The first synthetic glucocorticoids were developed in the late 1950s as drugs for treating rheumatoid arthritis. In basic research the compounds are mainly used to unravel the properties of the glucocorticoid receptor (GR).

All synthetic glucocorticoids are structurally related to hydrocortisone (cortisol) and corticosterone, which are the main endogenous glucocorticoids of primates and rodents, respectively. However, synthetic glucocorticoids often bind to the GR with a higher affinity and thus have much more potent glucocorticoid actions on the HPA axis, the immune system, and energy metabolism than the naturally occurring hormones. Their mineralocorticoid actions on salt retention/sodium homeostasis, although reduced, are still a disadvantage, especially for long-term treatment. The synthetic glucocorticoid RU28362 lacks this mineralocorticoid activity and is a pure glucocorticoid.

The group of glucocorticoids includes many compounds (see **Table 1** for some examples). The most commonly used synthetic glucocorticoids are prednisone/prednisolone and dexamethasone.

Chemical Aspects

The structure of all glucocorticoids is basically the same. Most of them are 21-carbon polycyclic compounds with a steroid structure, a fused 17-carbon atom ring system with an aliphatic side chain at carbon-17, a keto group at carbon-3, one or two double bonds in the A-ring, and an oxygen at carbon-11 (**Figure 1**). Different side groups determine the differences in glucocorticoid and mineralocorticoid properties. Reducing the salt-retaining properties was the main effort in the early development of new glucocorticoids. For example, the fluorination of dexamethasone (9α-fluoro-11β,17α,21-trihydroxy-16α-methylpregna-1,4-diene-3,20-dione) highly increases its anti-inflammatory potency, but unfortunately also its salt-retaining properties. The inserted 16α-methyl group, however, markedly reduces the mineralocorticoid properties, negating the effect of the 9α-fluoro

Table 1 Characteristics of several common glucocorticoids

Compound	*Relative affinity of glucocorticoid receptors (GRs)*	*Relative potency*		*Half-life (h)*		*Comments*
		Glucocorticoid	*Mineralocorticoid*	*Plasma*	*Biological*	
Hydrocortisone (cortisol)	1	1	1	1.5	8–12	Endogenous glucocorticoid of humans Drug of choice for replacement
Cortisone	0.01	0.8	0.8	0.5	8–12	Inactive until converted to hydrocortisone
Corticosterone	0.85	0.3–0.5	5–15		8–12	Endogenous glucocorticoid of rodents Not used in humans
Dexamethasone	7.1	25–30	0.5	3–4	36–72	Drug of choice for suppression of ACTH production
Prednisolone	2.2	3.5–4	(0.6–0.8)	2–3	12–36	
Prednisone	0.05	4	(0.6–0.8)	3–3.5	12–36	Inactive until converted to prednisolone
Methylprednisolone	11.9	5	0.5	3	12–36	
RU28362	9		0			Pure GR agonist; not used in humans
Betamethasone	5.4	25–30	0.5	6.6	36–54	
Triamcinolone	1.9	5	0.1	1.5	12–36	
Mifepristone (RU486)	7	–	–	20–30		GR antagonist

atom. Prednisolone (11β,17α,21-trihydroxypregna-1,4-diene-3,20-dione) is a dehydrogenated derivative of cortisol with an enhanced glucocorticoid and declined mineralocorticoid activity. The most selective ligand RU28362 (11β,17β-dihydroxy-6-methyl-17α (1-propynyl)-androsta-1,4,6-triene-3-one) has only been used in experimental research up to now. More recently developed selective GR ligands have a non-steroidal structure.

Pharmacokinetics

Glucocorticoids are small, highly lipophilic molecules and therefore are not very soluble in water, but are easily soluble in ethanol, polyethylene glycol (PEG), or dimethyl sulfoxide (DMSO).

Because of their strong lipophilicity, the orally administered glucocorticoids are readily absorbed. Prednisone and cortisone are biologically inactive and are quickly converted to the active compounds prednisolone and cortisol/hydrocortisone, respectively, by dehydrogenation at the carbon-11 hydroxyl group in the liver. In the blood, synthetic glucocorticoids are mainly bound to albumin, a common plasma protein with low affinity. Unlike the endogenous glucocorticoids, most synthetic glucocorticoids are not bound to corticosteroid-binding globulin (CBG). This plasma protein with a high affinity for naturally occurring corticosteroids, but low presence in plasma, binds prednisolone but not dexamethasone. Plasma half-life is enlarged in the presence of a double bond (prednisolone) or a fluorine atom (dexamethasone) because of decreased rate of metabolism. The plasma half-lives are between 2 and 4 h (Table 1). Because of the modifying action on gene transcription, the duration of the biological effects is much longer than the plasma half-life. It is not the presence of the drug itself but the long-lasting interaction of the ligand receptor complex and the protein synthesis system that accounts for the biological actions. The biological half-life can be up to 72 h (Table 1).

Synthetic glucocorticoids are metabolized in the same way as the endogenous corticosteroids. The main metabolizing tissue is the liver, where steroids are conjugated with glucuronic acid or sulfuric acid. The unchanged parent drug and its metabolites are secreted in the urine, for over 60% (in the case of dexamethasone) and up to 90% (in the case of prednisolone) of the administered dose within 24 h. Changes in plasma half-life of glucocorticoids may result from changed hepatic metabolism due to liver or renal malfunction.

The passage of synthetic glucocorticoids through the blood–brain barrier is hampered because of the presence of multidrug resistance 1a (mdr1a) P-glycoprotein. Recently it was shown that this drug efflux

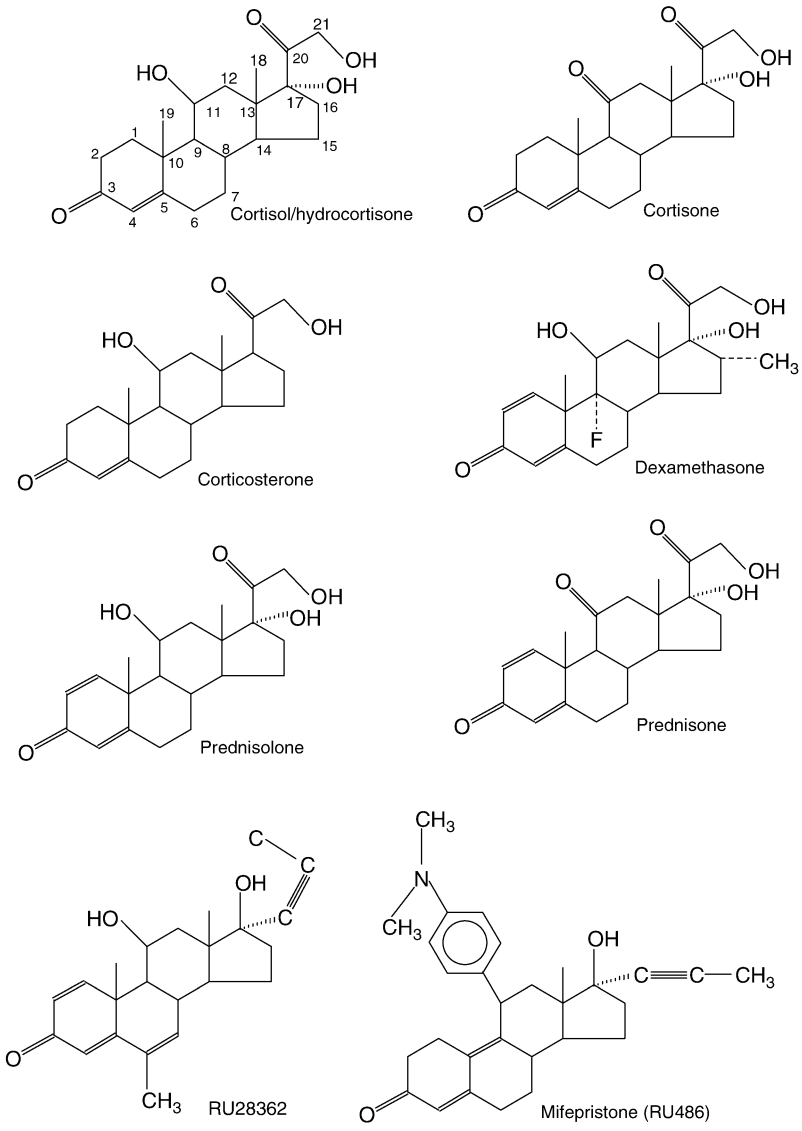


Figure 1 Chemical structures of some common glucocorticoids. Cortisol and corticosterone are endogenous corticosteroids. Cortisol and hydrocortisone have the same chemical structure; the former refers to the naturally occurring glucocorticoid, while the latter refers to the synthetic variant. Cortisone is an inactive endogenous derivative of cortisol. Prednisolone, prednisone, dexamethasone, and RU28362 are synthetic glucocorticoids. Mifepristone is a glucocorticoid antagonist.

pump, which is expressed at the apical membranes of endothelial cells of the blood–brain barrier, hampers the access of moderate amounts of dexamethasone and prednisolone to the brain. Ablation of the gene in the mdr1a knockout (−/−) mice resulted in enhanced uptake of dexamethasone and prednisolone.

The most effective receptor antagonist of glucocorticoids is mifepristone (RU486) [17β-hydroxy-11β-(4-dimethylaminophenyl)-17α-(1-propynyl)-estra-4, 9-dien-3-one]. This compound, originally developed as a progesterone antagonist, expresses a high affinity for the GR, but binding generally does not result in stimulation of transcription. RU486 antagonizes glucocorticoid effects *in vitro* and inhibits the effect of dexamethasone on HPA activity *in vivo*. RU486 has weak agonistic properties in some *in vitro* settings, depending on the molecular mechanism of action, and possibly in *in vivo* settings in the absence of glucocorticoid agonists.

Actions of Synthetic Glucocorticoids

The main glucocorticoid actions are mediated by GRs, nuclear receptors present in the cytosol of almost every cell in the body. The corticosteroids bind with high affinity to the GR and in some degree to mineralocorticoid receptors (MRs). MRs are less abundantly present in the body. Most of the synthetic

glucocorticoids have relatively low residual mineralocorticoid potencies (Table 1).

In absence of P-glycoprotein, glucocorticoids easily cross the cell membrane, entering the cell by passive diffusion. Upon binding, the activated steroid receptor complex is translocated from the cytoplasm to the nucleus. The next step is binding to DNA or to transcription factors. Binding to DNA involves binding to glucocorticoid responsive elements (GREs) and subsequent recruitment of coactivator proteins, resulting in activation (or inhibition) of the transcription of target genes (referred to as transactivation). Following binding to transcription factors, the steroid receptor complex effects the transcription of non-GRE containing genes (referred to as transrepression). Transrepression is believed to be the main mechanism of the anti-inflammatory/immunosuppressive actions of glucocorticoids, while transactivation is believed to be responsible for the numerous side effects. This concept has fueled the extensive search for GR ligands with reduced metabolic and HPA suppressant actions while retaining their immunosuppressive/anti-inflammatory actions. The first compounds with dissociated glucocorticoid activities at the molecular level did not show the desired dissociation of beneficial and undesirable actions *in vivo*. A new interesting approach is emerging in the literature that focuses on the development of improved glucocorticoids based on the concept of differential cofactor recruitment by the GR depending on the ligand involved. These new ligands are useful in the understanding of basic GR functioning.

Regarding stress, the main action of glucocorticoids is their suppression of stress-induced HPA axis activity via GRs. GRs are found in all HPA axis structures, hypothalamus, hippocampus, and pituitary. Synthetic glucocorticoids reduce basal plasma levels of ACTH and corticosterone/cortisol during the circadian rise, and they suppress stress-induced rises in ACTH and corticosteroid plasma levels. In spite of the high affinity for synthetic glucocorticoids displayed by GRs *in vitro*, synthetic glucocorticoids (e.g., dexamethasone) do not accumulate in the hippocampus, because of the presence of mdr1a P-glycoprotein efflux transporter at the blood–brain barrier. It is therefore likely that the pituitary gland is the primary site of action for the suppressant effects on stress-induced ACTH and corticosterone secretion, particularly at low doses of synthetic glucocorticoids.

Clinical Aspects

Synthetic glucocorticoids are therapeutically used mainly as effective and wide-acting anti-inflammatory and immunosuppressive drugs. They affect all types of inflammatory reactions independently of the initiating stimulus. They suppress the body's defense reactions, preventing overreaction of the defense mechanisms, which may become damaging if not controlled. Unfortunately, the wide range of actions of glucocorticoids also results in many, sometimes life-threatening side effects on bone, glucose, and HPA axis. More selective nonsteroidal glucocorticoid ligands are currently in development.

Another clinical application is substitution therapy in the case of adrenal insufficiency (Addison's disease), although cortisol and cortisone are more commonly used in that case. Dexamethasone is also used to reduce brain edema. For diagnostic purposes the dexamethasone suppression test (DST, sometimes combined with a CRH challenge) is widely used to test the hypothalamic-pituitary-adrenocortical function. A low dose results in suppression of the hydrocortisone secretion. Failure of suppression implies hypersecretion of corticotropin or of glucocorticoids due to Cushing's syndrome. In psychiatry the DST is used to test patients with major depressive disorders. A large proportion of depressed patients displays enhanced HPA activity. They show impaired suppressibility of cortisol and thus escape from dexamethasone suppression.

Mifepristone (RU486) was recently demonstrated to rapidly relieve psychotic depressed patients from psychotic symptoms. After a brief 7-day treatment with high doses of this antagonist, patients showed a dramatic improvement.

See Also the Following Articles

Corticosteroid Receptors; Corticosteroid-Binding Globulin (Transcortin); Corticosteroids and Stress; Cytokines; Dexamethasone Suppression Test (DST); Glucocorticoids, Overview; Immune Suppression.

Further Reading

Coghlan, M. J., Elmore, S. W., Kym, P. R. and Kort, M. E. (2003). The pursuit of differentiated ligands for the glucocorticoid receptor. *Current Topics in Medicinal Chemistry* **3**(14), 1617–1635.

De Kloet, E. R., Vreugdenhil, E., Oitzl, M. S. and Joëls, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews* **19**(3), 269–301.

Holsboer, F. and Barden, N. (1996). Antidepressants and hypothalamic-pituitary-adrenocortical regulation. *Endocrine Reviews* **17**(2), 187–205.

Miller, D. D., Brueggemeier, R. W. and Dalton, J. T. (2002). Adrenocorticoids. In: Foye, W. O., Lemke, T. L. &

Williams, D. A. (eds.) *Foye's principles of medicinal chemistry* (5th edn.), chap. 28. Philadelphia, PA: Lippincott Williams & Wilkins.

Munck, A., Guyre, P. M. and Holbrook, N. J. (1984). Physiological functions of glucocorticoids in stress and their relation to pharmacological actions. *Endocrine Reviews* 5(1), 25–44.

Schimmer, B. P. and Parker, K. L. (1996). Adrenocorticotropic hormone; adrenocortical steroids and their synthetic analogs; inhibitors of the synthesis and actions of adrenocortical hormones. In: Hardman, J. G. & Limbird, L. E. (eds.-in-chief) *Goodman and Gilman's the pharmacological basis of therapeutics* (10th edn.). chap. 60. New York, NJ: McGraw Hill.

Systemic Lupus Erythematosus

D J Wallace
Cedars-Sinai/David Geffen School of Medicine, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

Etiopathogenesis
Clinical Presentation
Management
Prognosis
The Role of Stress in the Immune Response
Clinical Implications of Stress in Lupus
Management of Stress in Systemic Lupus Erythematosus

Glossary

Antinuclear antibodies	Proteins in the blood that react with the nuclei of cells; seen in 96% of patients with SLE, most with other autoimmune diseases, and in 5% of healthy individuals, they are the hallmark of autoimmune disease.
Apoptosis	Programmed cell death; in SLE, the mechanism is altered such that cells that should die do not. Fragments from these cells lead to autoimmune reactions.
B cells	White blood cells that make autoantibodies.
Immune complexes	The complex of an antigen and antibody; in SLE, they are not cleared properly and damage tissue, producing an inflammatory response.
T cells	White blood cells responsible for immunological memory; in SLE, T suppressor cells fail to suppress immune responses.

Lupus erythematosus is an autoimmune disorder that afflicts 1 million people in the United States; 90% of patients are female, and 90% develop it during their reproductive years. Epidemiological surveys suggest that African Americans, certain Asian groupings (e.g., Chinese), Native Americans, and Hispanics are more susceptible to the disorder. Lupus is classified into many groupings: chronic cutaneous (discoid) lupus (15%), systemic lupus (75%), drug-induced lupus (5%), and lupus overlapping with rheumatoid arthritis, scleroderma, or inflammatory myositis (also called cross-over syndromes, mixed connective tissue disease) (5%).

Etiopathogenesis

A critical dose of 30 thus far identified susceptibility genes causes enough immune response abnormalities to sustain the production of pathogenic autoantibodies and immune complexes. The genes are turned on by infections, increased exposure to ultraviolet light or other environmental toxins, at least 70 known prescription drugs, exposure to increased estrogen or other sex hormones, and severe stress. Antigens are formed that are immunogenic, lead to high quantities of apoptotic cells, produce molecular mimicry (especially microbes), and/or undergo epitope spreading. There are intrinsic abnormalities of B and T lymphocytes. These cells are activated by lower concentrations of antigens than normal, have sustained surface expression of activation markers, and are relatively resistant to apoptosis, which allows them to escape regulation. There is a consequent release of cytokines and chemokines by these cells, which results in inflammation. In this milieu, autoantibody subsets become pathogenic and idiotypic networks promote more antibody production. Immune complexes (consisting of antigen and antibody) are formed that favor capture rather than elimination, and tissue is further damaged. The role that stress plays in this process is discussed here.

Clinical Presentation

Approximately one-half of all systemic lupus erythematosus (SLE) cases are non-organ threatening.

In other words, the heart, lung, kidney, liver, central nervous system, and hematopoietic system (with regards to hemolytic anemia and autoimmune thrombocytopenia) are spared. Patients usually complain of constitutional symptoms of fatigue, fevers, swollen glands, or weight less along with arthralgias, myalgias, and pleuropericardial (serositis) discomfort. These individuals are usually diagnosed after 1–3 years of symptoms. Organ-threatening disease is a serious subset of lupus and is associated with the need for immune-suppressive approaches. Manifestations include myocarditis, endocarditis, pulmonary hypertension or hemorrhage, mesenteric vasculitis, interstitial lung disease, autoimmune hepatitis, thrombotic thrombocytopenic purpura, retinitis, vestibulitis with hearing loss, proteinuria, seizures, psychosis, and edema. The presentation is usually quite dramatic and diagnostic delays are uncommon. The 1997 American College of Rheumatology (ACR) revised criteria for SLE are listed in Table 1.

Laboratory testing, along with appropriate electrical studies and imaging, is very useful in diagnosing and following the disorder. Initial testing includes a complete blood count, comprehensive metabolic profile, muscle enzymes, urinalysis, chest x-ray, and electrocardiogram. Serologic parameters encompass an antinuclear antibody (ANA) with a reflex panel for a variety of autoantibodies, obtaining certain protein and acute-phase reactant determinants (which screen for inflammation) such as anti-DNA, anti-Smith antibodies (anti-Sm), Sjögren's syndrome antigens (anti-SSA and anti-SSB), complement components, sedimentation rate, C reactive protein, and a serum protein electrophoretic screening. Occasionally, additional determinations such as two-dimensional echocardiography, brain imaging, lumbar puncture, or electromyograms can be useful. One-third of lupus patients have a procoagulant set of autoantibodies known as antiphospholipid antibodies that can predispose patients to a thromboembolic event.

Management

Non-organ-threatening disease is managed by employing physical measures such as avoidance of the sun, pacing activities to minimize fatigue, moist heat for painful joints, a general conditioning program, work-station modifications to enhance job performance, and a low fat, well-balanced diet. Stress reduction is also critical. Individuals respond best to a combination of sunscreens applied to sun-exposed areas, topical corticosteroids, and nonsteroidal anti-inflammatories. Antimalarials (e.g., hydroxychloroquine) have anti-inflammatory disease-modifying properties and can decrease the risk of dissemination. Low doses of corticosteroids may be used until antimalarials are effective several months later, and immune-suppressive regimens are rarely indicated.

Patients with organ-threatening disease are usually managed with a combination of corticosteroids (moderate to high dose, usually starting at 1 mg/kg/day prednisone or its equivalent). Steroid-sparing agents such as methotrexate and leflunomide can be employed for milder cases, but serious presentations respond best to azathioprine, mycophenolate mofetil, or cyclophosphamide. A variety of biological regimens and stem cell options are currently being explored.

Prognosis

Patients with non-organ-threatening disease who do not have antiphospholipid antibodies have a normal life expectancy. Organ involvement increases morbidity and mortality. Approximately 50% have 20-year life expectancies. The most common causes of demise

Table 1 Revised American College of Rheumatology criteria for SLE (1997)[a]

Cutaneous criteria	Butterfly rash (lupus rash over the cheeks and nose)
	Chronic cutaneous (discoid) rash (a thick, plaquelike, adherent rash that scars)
	Sun sensitivity (rash after being exposed to ultraviolet A or B light)
	Oral ulcerations (recurrent sores in the mouth or nose)
Systemic criteria	Arthritis (inflammation of two peripheral joints with tenderness, swelling, or fluid)
	Serositis (inflammation of the lining of the pleura or pericardium)
	Neurological disorder (seizures or psychosis with no other explanation)
Laboratory criteria	Blood abnormalities (hemolytic anemia, leucopenia, thrombocytopenia)
	Immunological disorder (antiphospholipid antibodies, anti-dsDNA, anti-Sm, lupus anticoagulant or false-positive syphilis serology)
	Positive antinuclear antibody blood test

[a]A person is said to have SLE if four of the 11 criteria are present at any time. Anti-dsDNA, anti-double-stranded DNA antibodies; anti-Sm, anti-Smith antibodies.

include infection, complications from accelerated atherogenesis and immune suppression, and unresponsive inflammation. For those with ongoing chronic disease, quality of life can be markedly diminished due to medication (e.g., corticosteroids), swollen joints and inflammation limiting activities of daily living, chronic pain, specific diets, sun avoidance, decreased ability to think clearly, edema from nephrosis, and cold avoidance due to Raynaud's disease, among other factors. This is associated with increased stress and decreased ability to respond to it.

The Role of Stress in the Immune Response

The role of stress in SLE has been looked at in a small number of animal models and human studies. Stress is a complex concept that involves interactions of the hypothalamic-pituitary-adrenal (HPA) axis in combination with cytokines, hormones, and the sympathetic nervous system. Patients with SLE generally manifest a blunted HPA axis response and relative glucocorticoid resistance. The latter has been documented by the larger doses needed to achieve a therapeutic effect in managing lupus, smaller numbers of glucocorticoid receptors on cells, and defectively functioning or mutated glucocorticoid receptors. Chronic inflammation is associated with hypercortisolism. In addition, prolactin is an established acute-phase reactant. Levels are increased in up to 20% of lupus patients, and this hormone can also increase interferon and interleukin (IL-)2 production.

Recently, there has been considerable interest in the role of cytokines in SLE. Lupus mice have diminished IL-1 receptor expression, and lupus patients have increased serum IL-6 levels that correlate with flares. Cerebrospinal fluid levels of interferon α are elevated in lupus psychosis. The release of these glycoproteins is a well-documented stress response. The release of cortisol is also associated with increased sympathetic nervous system activity. The autonomic nervous system is dysfunctional in lupus, as well established by Raynaud's phenomenon and vasomotor instability (e.g., palmar erythema). This has been correlated in the disorder with the increased release of neuropeptide Y.

In summary, lupus patients have a sluggish corticotropin releasing hormone (CRH) response, relative glucocorticoid resistance, and hypercortisolism suggestive of an adaptive response to chronic inflammation. Cytokine and hormonal imbalances along with dysautonomia serve to perpetuate this.

Clinical Implications of Stress in Lupus

Lupus is associated with substantial emotional and societal costs. It has been suggested that up to half of all lupus patients have psychiatric comorbidities such as dysthymia and depression. This section reviews whether stress can induce or exacerbate lupus and how it can be managed (Table 2).

Various studies have shown that emotional stress and physical trauma can affect the immune system by causing decreased lymphocyte mitogenic responsiveness, lymphocyte cytotoxicity, increased natural killer cell activity, skin homograft rejection, graft-versus-host response delayed hypersensitivity. Could the impairment in T-cell immune functions be related to a clinical flare mediated by B-cell hyperreactivity? Acute stress in SLE has been purported to correlate with increased urine neopterin levels, IL-4 levels, decreased natural killer cell response, and increased β_2-adrenoreceptors on mononuclear cells without a clinical disease flare.

Can Emotional Stress Induce Lupus?

As early as 1955, crises in interpersonal relationships were associated with the onset of disease. Up to one-half of all lupus patients interviewed believe that stress is associated with its initial presentation. Unfortunately, no evidence-based study has documented

Table 2 Stress in SLE – summary[a]

Sluggish CRH response, relative glucocorticoid resistance, and hypercortisolism suggestive of an adaptive response to chronic inflammation
Cytokine abnormalities, hormonal dysfunction, and dysautonomia perpetuate this response
Physical trauma can produce cutaneous lupus; no evidence-based studies have shown that physical or emotional trauma can induce SLE
Evidence-based studies have documented that physical trauma and emotional stress can aggravate SLE, although concomitant fibromyalgia is confounding factor and strict guidelines of abnormalities occurring within 60 days of a specific incident have been published for litigation purposes
Stress-reduction measures include antidepressants, anxiolytic agents, biofeedback, cognitive-behavioral therapy, and counseling.

[a]CRH, corticotropin releasing hormone.

this. Patients with SLE possess autoantibodies several years prior to diagnosis, according to a serum repository kept by the U.S. Armed Forces. Many patients manifest an undifferentiated connective tissue disease (UCTD) for years prior to evolving into frank SLE. (UCTD is defined as inflammatory arthritis or Raynaud's with a positive ANA or rheumatoid factor in an individual who fails to meet the ACR criteria for lupus or rheumatoid arthritis. One-third of patients evolve SLE or rheumatoid arthritis, one-third remain UCTD, and one-third have spontaneous resolution of their symptoms.) The mechanisms resulting in this transformation are not known.

Can Emotional Stress Exacerbate Lupus?

At Harvard University, of 45 flares in 160 patients over a 40-year period, 41 patients associated these exacerbations to emotional stressors. Lupus patients may have significantly reduced cell mobilization to psychological stress compared to controls. Several surveys have subsequently shown that stress can elevate antinuclear antibody levels and be associated with clinical flares. Life-event stress predicted adverse changes in functional disability over a 6-month period in one study and self-reported daily stressors were associated with increases in clinical activity in three others. These studies did not take into account the impact of quality of life related to steroid therapy for lupus or the confounding influences of fibromyalgia, which is present in one-third of all lupus patients.

Can Physical Trauma Induce or Exacerbate Lupus?

Cutaneous skin lesions seen in lupus may be found after physical trauma, lacerations, and scars. Sometimes these scars evolve into skin cancers. Two studies, on the other hand, suggested that major physical trauma such as an earthquake is not associated with disease flares but may actually improve the process. The physical trauma of a motor vehicle accident is also associated with myofascial injuries, and no published study has associated this with a lupus flare. Guidelines for documenting a lupus flare after trauma have been published and include the following findings within 60 days of the event: (1) increased symptoms and physical signs of lupus documented in a progress note, (2) worsening of inflammatory laboratory parameters (e.g, sedimentation rate or anti-DNA), (3) increases in lupus medication or new anti-inflammatory medication ordered by the practitioner, and (4) introduction of psychotropic medication prescribed by the treating physician or referral to a psychologist documenting increased stressors. The patients medical record over the prior year can be used as a historical control.

Management of Stress in Systemic Lupus Erythematosus

Stress-reduction measures include managing the inflammation associated with disease activity, counseling, pain management, cognitive-behavioral therapy, and anxiolytics and antidepressant medications. Counseling usually emphasizes improved coping, cognitive therapy, and biofeedback; improving resilience; managing stress, and regulating mood. Controlled studies have validated this approach.

Further Reading

Adams, S. G., Jr, Dammers, P. M., Sala, T. L., et al. (1994). Stress, depression and anxiety predict average symptom severity and daily symptom fluctuation in systemic lupus erythematosus. *Journal of Behavioral Medicine* **17**, 459–477.

De Bellis, M. D., Burke, L. and Trickett, P. K. (1996). Antinuclear antibodies and thyroid function in sexually abused girls. *Journal of Traumatic Stress* **9**, 369–378.

Dobkin, P. L., Fortin, P. R., Joseph, L., et al. (1998). Psychosocial contributors to mental and physical health in patients with systemic lupus erythematosus. *Arthritis Care Research* **11**, 23–31.

Eskreis, B. D., Eng, A. M. and Furey, N. L. (1988). Surgical excision of trauma-induced verrucous lupus erythematosus. *Journal of Dermatologic Surgery and Oncology* **14**, 1296–1299.

Gorby, H. E. and Sternberg, E. M. (2007). Neuroendocrine immune interactions: principles and relevance to systemic lupus erythematosus. In: Wallace, D. J. & Hahn, B. H. (eds.) *Dubois' lupus erythematosus*, pp. 286–300 (7th edn.). Philadelphia: Lippincott Williams & Wilkins.

Greco, C. M., Rudy, T. E. and Manzi, S. (2004). Effects of a stress-reduction program on psychological function, pain, and physical function of systemic lupus erythematosus patients: a randomized controlled trial. *Arthritis & Rheumatism* **51**, 625–634.

Jacobs, R., Pawlak, C. R., Mikeska, E., et al. (2001). Systemic lupus erythematosus and rheumatoid arthritis patients differ from healthy controls in their cytokine pattern after stress exposure. *Rheumatology* **40**, 868–875.

King-Smith, D. (1926). External irritation as a factor in the causation of lupus erythematosus discoides. *Archives of Dermatology* **14**, 547–549.

McLeary, A. R., Meyer, E. and Weitzman, E. L. (1955). Observations on the role of depression in some patients with disseminated lupus erythematosus. *Psychosomatic Medicine* **17**, 311–321.

Pawlak, C. R., Witte, T., Heiken, H., et al. (2003). Flares in patients with systemic lupus erythematosus are associated with daily psychologic stress. *Psychotherapy and Psychosomatics* **72**, 159–165.

Peralta-Ramirez, M. I., Jimenez-Alonso, J., Godoy-Garcia, J. F., et al. (2004). The effects of daily stress and stressful life events on the clinical symptomatology of patients

with lupus erythematosus. *Psychosomatic Medicine* **66**, 788–794.

Ropes, M. W. (1976). *Systemic lupus erythematosus*. Cambridge, MA: Harvard University Press.

Wallace, D. J. (1994). Does stress or trauma aggravate rheumatic disease? *Baillière's Clinical Rheumatology* **8**, 149–159.

Wallace, D. J. and Hahn, B. H. (eds.) (2007). *Dubois lupus erythematosus* (7th edn.). Philadelphia: Lippincott, Williams & Wilkins.

Wallace, D. J. and Metzger, A. L. (1994). Can an earthquake cause flares of rheumatoid arthritis or lupus nephritis? *Arthritis & Rheumatism* **37**, 1826–1828.

T

T Cells *See:* Lymphocytes.

Teaching and Stress

E R Greenglass
York University, North York, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E R Greenglass, volume 3, pp 571–574, © 2000, Elsevier Inc.

Stressors in Teaching
Teacher Burnout
Burnout as a Process

Glossary

Burnout	A state of physical, emotional, and mental exhaustion that results from long-term involvement with people in situations that are emotionally demanding.
Emotional exhaustion	Feelings of being emotionally overextended and drained by others.
Depersonalization	A callous response toward people who are the recipient of one's services.
Cynicism	An indifferent or distant attitude towards one's work.
Lack of personal accomplishment	A decline in one's feelings of competence and successful achievement in one's work with people.
Social support	The exchange of information leading a person to believe that he or she is cared for. It can also involve provision of information, tangible help, or emotional help.
Role conflict	The interference of the efforts to fulfill the demands associated with one role with the efforts to fulfill the demands associated with other roles.

Stressors in Teaching

It is well documented that teaching is a highly stressful occupation. Several writers have noted recurring demands in the teaching environment that lead to significant levels of stress and burnout in teachers. There are several sources of stress for teachers. One of the most frequently acknowledged stressors for teachers centers around students. Problems with students include disruptive behavior, such as verbal and physical abuse; emotional demands of students; their special needs; and their heterogeneous in abilities – all of which tax the time and energy of an already busy teacher. Another major stressor for teachers involves the conflicting demands made by supervisors, colleagues, students, and the students' parents. A third set of stressors in teaching relates to teachers' workload itself, mainly their work overload. Quantitative work overload involves too many demands on teachers with too little time in which to meet these demands adequately. Often it is accompanied by feelings of being rushed and impatience. The importance of work overload as a stressor is underlined by the fact that teachers themselves in countries all over the world have consistently cited work overload as a major stressor in their job. Important dimensions associated with work overload are excessive paperwork, lack of preparation time for class, making reports, and submitting grades. Having to meet constant deadlines throughout the year, including those associated with grading, tests, exams, and written reports by students is cited as another major stressor by teachers. Oversized classes with either inadequate or no teaching assistance also function as stressors in teachers.

Additional stressors associated with teaching are found at the administrative level. A frequently cited administrative stressor is the exclusion by the administration of teachers from decision making that bears directly on their workload and the quality of their work life. Research has shown that when teachers do not have the opportunity to participate in the school's decision-making processes, their morale, motivation, self-esteem, and job satisfaction are more likely to decline. Lack of recognition of teachers' professional accomplishments by supervisors and school administrators alike contributes to an increase in stress levels in teachers. When teachers experience high role conflict in their jobs, role ambiguity, and low autonomy, they are more likely to experience stress. At the administrative level, additional stressors include inflexible employment policies on staffing and leave, and poor communication with teachers. Thus, research has demonstrated that teaching is associated with a wide range of stressors, including those intrinsic to teaching, those involving the relationship between teacher and student, and those related to the organizational structure of the school including the school administration.

Teacher Burnout

Considerable research worldwide has examined psychological consequences of stress in teachers. Over time, the cumulative effects of stressors lead to a decrease in job satisfaction, reduced commitment to teaching, absenteeism, and job turnover. At the individual level, greater stress is associated with an increase in somatic complaints and with physical and psychological illness. Research has shown that the experience of stress over time often leads to burnout in teachers. Although burnout is linked to occupational stress, burnout goes beyond stress by emphasizing the accumulation of the effects of stressors within a context that emphasizes total life and environmental experiences. Research findings indicate that teacher burnout in several Western industrialized countries has increased and affects a significant and growing proportion of teachers. Burnout is often defined as a state of physical, emotional, and mental exhaustion that results from long-term involvement with people in situations that are emotionally demanding. Because individuals who enter teaching are often presented with work pressures in addition to teaching, they often find themselves in situations that are heavily demanding. Research shows that human service professionals, including teachers, are particularly vulnerable to burnout. Freudenberger's original idea of burnout was that workers would find themselves under increasing pressure to help others, would demand more of themselves than they were able to give, and would ultimately exhaust themselves.

Measurement of Burnout

Given the significance of burnout for the individual, it is necessary to have a standardized instrument to measure it. The Maslach Burnout Inventory (MBI), developed by Christine Maslach and Susan Jackson in 1986, is currently the research instrument used most widely to measure burnout. The MBI has three subscales that measure three different aspects of burnout. These are emotional exhaustion, depersonalization, and reduced personal accomplishment. Emotional exhaustion is defined as feelings of being emotionally overextended and drained by others. Depersonalization is a callous response toward people who are the recipient of one's services.

Lack of personal accomplishment is a decline in one's feelings of competence and successful achievement in one's work with people. The need for developing a scale that measures burnout in occupational groups other than human service professionals prompted the development of the Maslach Burnout Inventory-General Survey (MBI-GS) in 1996. Thus, the MBI-GS defines burnout as a crisis in one's relationship with work, not necessarily as a crisis in one's relationship with people. In the MBI-GS, the depersonalization scale is replaced with a scale measuring cynicism, an indifferent or distant attitude toward one's work. Research findings show that teachers are vulnerable to burnout in all three areas. For example, teachers show emotional exhaustion when they perceive they are unable to give of themselves to students as perhaps they did earlier in their careers. Depersonalization and cynicism occur when teachers develop negative, cynical, and sometimes callous attitudes toward students, parents, and/or colleagues. Feelings of diminished personal accomplishment occur when teachers perceive themselves as ineffective in helping students to learn and in fulfilling other school responsibilities. For individuals who experience burnout, one of the first reactions is distancing themselves from the job. Burnout has also been linked to job turnover, absenteeism, poor performance, and various types of individual distress symptoms including depression, anxiety, and somatization. Overall, burned-out teachers are likely to be less sympathetic toward students, have a lower tolerance for classroom disruption, be less apt to prepare adequately for class, and feel less committed to their work.

Gender Differences and Burnout

One question that has been raised by many researchers regarding teachers is: who is more burned out, women or men? Research findings from a variety of studies conducted in several different countries show that men have significantly greater levels of depersonalization than women. Why should men be more prone to depersonalization, an attitude characterized by callousness and being impersonal toward one's students? One explanation is found in accepted norms associated with the masculine gender role, which emphasizes strength, independence, separation, and invulnerability. In this context, depersonalization may be regarded as a reflection of men's repressed emotionality. Described in this way, depersonalization may be seen as an ineffective coping form that allows men to continue their work with people yet remain untouched in any significant way by others' suffering. Moreover, it is through the distance they have imposed between their students and themselves that men can remain uninvolved, thus appearing callous. It is also possible that depersonalization is a result of an inability on the part of men to cope with work strain. Because it is socially unacceptable for men to openly express vulnerability, men have fewer options for emotional expression. Higher depersonalization in men has been associated with lack of collegiality, higher absenteeism, greater medication use, and a lower quality of life.

Another explanation for women's lower depersonalization compared to men focuses on women's greater ability to cope with interpersonal stress. Women have an edge over men in dealing with the emotional strain of people work because the feminine gender role emphasizes caring, nurturance, and concern for others. As a result, women may be less likely than men to respond to troubled individuals in an impersonal and callous manner. Some researchers have found that women have higher scores than men on emotional exhaustion. Women may experience greater emotional exhaustion because, compared to men, their total workload (including paid and unpaid work) is higher. Although the number of men doing housework and childcare has increased in recent years, according to research findings women still do most of the work in the home. Further research has shown that women's increased workload interferes with their ability to wind down from a stressful day, with resultant negative effects on their health. Research also shows that women experience greater role conflict (between work and family roles) than their male counterparts. Related findings indicate that there are different predictors of burnout in men and women. In women role conflict between home and work roles is a significant predictor of burnout, whereas in men the predictors of burnout tend to be confined to work. A structural explanation of gender-related differences in burnout highlights the association of role conflict with stress and burnout in women. It is widely recognized that employed women, including teachers, typically maintain major responsibility for the home and the family, and thus more often than men, bear the burden of role conflict.

Burnout, Stress, Social Support and Coping

Over the years, researchers have investigated several ways that burnout in teachers can be reduced, including individual and social forms. Many of the social forms involve interventions that have led to lower burnout levels in teachers by changing the work environment or reducing burnout at its source (i.e., in the workplace). Specific examples include teacher involvement and participation in the decision making undertaken by the school; facilitating the development and provision of social support; and improving supervision through the clarification of work goals, which has resulted in less role ambiguity and lower role conflict in teachers. Burnout can be reduced by establishing clear lines of authority and responsibility in helping to reduce ambiguity and conflict. With teachers, it may be valuable to solicit their input in areas that affect their work life in order to gain greater involvement and commitment. The development of social support among teachers can be achieved by providing adequate time and location interaction. Encouraging the development of mentoring relationships between older and younger teachers can also be useful in increasing commitment among teachers. Considerable research has demonstrated the effectiveness of social support in lowering burnout and stress levels in teachers. Social support is the exchange of information leading a person to believe that he or she is cared for. It can also involve the provision of information and tangible, practical, and/or emotional help. Research suggests that social support may be effective in reducing burnout directly or as a buffering agent. When there is a direct effect, social support is positively related to physical and psychological health, regardless of the presence or absence of work stressors. Social support may also moderate the impact of stress on burnout so as to assist people with high stress to cope better. The buffering argument suggests that stress may affect some people adversely but that those with social support resources are relatively resistant to the deleterious effects of stressful events.

Several studies report that teachers with high burnout perceive less social support in their school

environment than those with lower burnout levels. Teachers' perceptions of stress are reduced by perceived support, particularly from school staff, including colleagues. Further research findings indicate that teachers who work in an environment that they perceive as being supportive are less likely to experience high levels of stress and burnout. Research by Greenglass and co-workers indicated that support from a teacher's peers or coworkers are the most important buffers of burnout. As suggested by Pines, social support in the form of technical advice such as helping teachers find more successful and innovative ways of working with students, may be a particularly effective intervention for burned-out teachers. Coping may also lead to lower levels of burnout. For example, in a 2002 study of burnout in German teachers, the relationship between proactive coping and burnout was examined. In this study, 316 German teachers were surveyed and job burnout was defined three-dimensionally in terms of emotional exhaustion, depersonalization, and lack of personal accomplishment. Proactive coping is multidimensional and forward-looking. It combines autonomous goal setting with self-regulatory goal attainment cognitions. Proactive coping integrates processes of personal quality of life management with those of self-regulatory goal attainment. The results showed a significant pattern of decreasing burnout with increasing levels of proactive coping. High-proactive-coping teachers reported less emotional exhaustion, less depersonalization, and more personal accomplishment than low-proactive-coping teachers.

Burnout as a Process

Increasingly, research has focused on the examination of burnout as a process that develops over time. As such, researchers have been interested in discovering the paths through which burnout develops. Some researchers have demonstrated the existence of two separate paths to burnout: a cognitive path, manifested in personal and professional feelings of lack of accomplishment, and an emotional path, reflected in a sense of overload and emotional exhaustion. It is generally held that the process begins to develop with external stressors, such as disruptive students, excessive paper work, and conflicting demands, which lead to emotional exhaustion. Emotional exhaustion is considered the aspect of burnout that is most responsive to the nature and intensity of work stressors. Some researchers refer to emotional exhaustion as the prototype of stress. In general, various job conditions are more strongly related to emotional exhaustion than to the other two subscales, depersonalization and lack of personal accomplishment. Emotional exhaustion is predicted mainly by occupational stressors, such as work overload, inadequate use of skills, and interpersonal conflicts. In addition, emotional exhaustion is considered the affective component of burnout that leads to depersonalization or cynicism. Through depersonalization, individual teachers attempt to cease the depletion of their emotional energy by treating their students as objects rather than people. Emotional exhaustion can also lead to reduced feelings of personal accomplishment in the teaching role, which may be seen as indicative of an impoverishment of people's perceptions of themselves in their work role, particularly as teachers. Findings suggest that depersonalization contributes to lower feelings of accomplishment as well. Because effective teaching is contingent on teachers' communicating with students on a personal level, teachers may be particularly vulnerable to low feelings of accomplishment when perceiving that they are treating their students impersonally, as in the process of depersonalization. Other findings point to the consistency of each of the burnout components over time. Emotional exhaustion, depersonalization, and lack of personal accomplishment predict significantly the scores on the same scales 1 year later, according to Greenglass and co-workers. Researchers agree that burnout levels are fairly stable over time.

One of the advantages of viewing burnout as a process over time is that it allows the investigation of the antecedents of burnout, particularly those associated with the school, the teachers' interpersonal relationships, and their workload. Being able to specify the antecedents of burnout in teachers has both theoretical and applied value. By specifying variables that contribute to burnout, theoretical knowledge of the process of the development of stress and burnout is advanced. On a practical level, it is valuable to the school's administration to be able to specify just what causes burnout. Examining the development of stress and burnout over time offers the possibilities of changing the processes involved and grappling with the complex causal relationships. This presumably can lead to the development of better and more effective intervention techniques, which can improve teacher morale and prevent the harmful consequences of further burnout.

Further Reading

Burke, R. J. and Greenglass, E. (1995). A longitudinal study of psychological burnout in teachers. *Human Relations* **48**, 187–202.

Freudenberger, H. (1974). Staff burnout. *Journal of Social Issues* **30**, 159–164.
Greenglass, E. R. (1991). Burnout and gender: theoretical and organizational implications. *Canadian Psychology* **32**, 562–570.
Greenglass, E. R. (2002). Proactive coping. In: Frydenberg, E. (ed.) *Beyond coping: meeting goals, vision, and challenges*, pp. 37–62. London: Oxford University Press.
Greenglass, E. R., Burke, R. J. and Konarski, R. (1997). The impact of social support on the development of burnout in teachers: Examination of a model. *Work and Stress* **11**, 267–278.
Greenglass, E. R., Fiksenbaum, L. and Burke, R. J. (1996). Components of social support, buffering effects and burnout: implications for psychological functioning. *Anxiety, Stress and Coping* **9**, 185–197.
Maslach, C. and Jackson, S. E. (1986). *Maslach Burnout Inventory manual* (2nd edn.). Palo Alto, CA: Consulting Psychologists Press.
O'Laughlin, E. M. and Bischoff, L. G. (2005). Balancing parenthood and academia: work/family stress as influenced by gender and tenure status. *Journal of Family Issues* **26**, 79–106.
Schaufeli, W. B., Leiter, M. P., Maslach, C., et al. (1996). The Maslach Burnout Inventory (MBI-GS). In: Maslach, C., Jackson, S. E. & Leiter, M. P. (eds.) *MBI manual* (3rd edn.). Palo Alto, CA: Consulting Psychologists Press.
Schwarzer, R. and Taubert, S. (2002). Tenacious goal pursuits and striving toward personal growth: proactive coping. In: Frydenberg, E. (ed.) *Beyond coping: meeting goals, visions and challenges*, pp. 19–35. London: Oxford University Press.

Relevant Websites

Greenglass, E. R., Schwarzer, R. and Taubert, S. (1999). The Proactive Coping Inventory (PCI): a multidimensional research instrument. http://www.psych.yorku.ca

Temperature Effects

J Roth
University of Giessen, Giessen, Germany

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J Roth, volume 3, pp 575–576, © 2000, Elsevier Inc.

Glossary

Heterothermy/ poikilothermy	The pattern of temperature regulation in a species in which the variation in core temperature, either nychthemerally or seasonally, exceeds that which defines homeothermy.
Homeothermy	The pattern of temperature regulation of a given species in which the cyclic variation in core temperature, either nychthemerally or seasonally, is maintained within arbitrarily defined limits despite much larger variations in ambient temperature.
Q_{10}	The ratio of the rate of a physiological process at a particular temperature to the rate at a temperature 10°C lower, when the logarithm of the rate is an approximately linear function of temperature.

Temperature change has striking effects on many physiological processes. All chemical reactions in an organism, including the processes of metabolism, are temperature-dependent. Within certain limits, an increase of temperature accelerates most physiological processes, such as the rate of oxygen consumption, which can be regarded as an expression for the overall metabolic activity of an organism. The rate of a physiological process usually increases with temperature in a regular manner. In oxygen consumption, for example, a rise of 10°C in temperature may cause a twofold or even threefold increase. The increase in the rate of a metabolic process caused by a 10°C increase in temperature is called the Q_{10} of this process. **Figure 1** shows an example for a given physiological process with a Q_{10} of 2. For a 10°C interval, the Q_{10} is thus defined as the ratio of the rate of the metabolic process ($R1$) at a particular temperature ($T + 10$) to the rate ($R2$) at a temperature 10°C lower (T), when the logarithm of the rate is an approximately linear function of temperature.

$$Q_{10} = R1_{(T+10)}/R2_{(T)}$$

According to van't Hoff's rule, the Q_{10} for thermochemical reactions is between 2 and 3, and so are the Q_{10} values for metabolic processes in an organism.

Animals can be divided into two groups: poikilotherms and homeotherms. The term poikilotherm

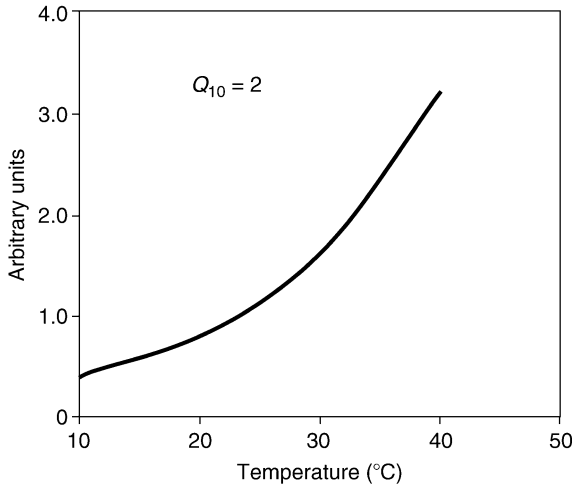


Figure 1 Temperature dependence of a given metabolic process. For this process, a temperature increase of 10°C results in the doubling of the rate of the process ($Q_{10} = 2$).

means that an animal shows a large variability of body temperature as a function of ambient temperature. In homeothermic animals (birds and mammals, including humans), the variations in body temperature are maintained within narrow limits despite much larger fluctuations in ambient temperature. This means that in poikilotherms the rate of metabolic processes is strongly influenced by temperature changes according to van't Hoff's rule. Van't Hoff's rule applies to homeotherms in the same way, but it is masked by other effects. The direct effects of changing temperatures on the metabolic processes within homeotherms are prevented by temperature regulation, the prerequisite to preventing large fluctuations of body temperature. As a consequence, the vital processes of poikilotherms depend entirely on external temperature conditions, whereas homeotherms live at practically the same level of intensity during the entire year and in all climatic regions. The geographic distribution of homeotherms is thus not as limited as the distribution of poikilotherms. Note that, in reality, the distinction between homeothermy and poikilothermy is not as clear-cut as the definitions of these terms suggest. On the one hand, many poikilothermic animals show rudimentary signs of temperature regulation, especially by behavioral means and, to a certain extent, even by metabolic responses. On the other hand, a number of homeothermic animals sometimes permit a profound fluctuation of body temperature (e.g., hibernators), with strong temporary restrictions of their vital processes.

In spite of their ability to regulate body temperature, large fluctuations in ambient temperature represent stressful conditions for homeotherms, especially when people or animals are exposed to temperatures that are out of their thermoneutral zone. For humans, the thermoneutral zone, or zone of thermal indifference, is defined as the range of ambient temperatures associated with specific water vapor pressure, air velocity, and radiant exchange within which 80% of active people do not complain of the thermal environment. A number of behavioral, physiological (functional), and morphological strategies with regard to the thermoregulatory system have developed to reduce the strain that is imposed on homeotherms by pronounced external temperature changes. These strategies can be summarized as thermal adaptations. Morphological adaptations, such as changes in body shape and insulation, need from months to years to develop unless they are genetically fixed and appear seasonally. In general, they are preceded by functional (physiological) modifications, including changes in the capacity of thermoregulatory effector systems and regulatory characteristics that need much less time to develop. One of the best-investigated changes in the capacity of thermoregulatory effector systems is the development of nonshivering thermogenesis in the brown adipose tissue of small mammals during cold adaptation. An increase in maximal sweat rate, and thereby in the capacity to dissipate heat, during adaptation to heat has also been described. The development of increased capacities of heat production and heat dissipation takes place within several weeks. Short-term physiological adaptations to stressful external temperatures consist of changes in thermoregulatory characteristics, which can be defined as deviations in the threshold and the gaining of the thermoregulatory responses. For example, a short exposure to a cold environment evokes not only thermoregulatory reflexes and emergency reactions but also a temporary shift of the shivering threshold to a lower mean body temperature. Continuous confrontation with a cold environment for several weeks results in an even more pronounced shift of the threshold for heat production to a lower body temperature. Similar adaptive changes of thermoregulatory characteristics occur in response to short- or long-term exposure to heat.

See Also the Following Article

Critical Thermal Limits.

Further Reading

Cossins, A. R. and Bowler, K. (1987). *Temperature biology of animals*. London: Chapman and Hall.

Fregly, M. J. and Blatteis, C. M. (1996). *Handbook of physiology. Vol. 1: Environmental physiology*. New York: Oxford University Press.

Precht, H., Christophersen, J., Hensel, H. and Larcher, W. (1973). *Temperature and life*. Berlin: Springer.

Terrorism

P Bell
South and East Belfast Health and Social Services Trust, Belfast, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Bell, volume 3, pp 577–580, © 2000, Elsevier Inc.

Introduction
Effects of Terrorism on the Terrorist
Effects of Terrorism on Victims
The Effect of Terrorism on Society

Glossary

Anomie	The sense of isolation that comes from lack of social integration and cohesion.
Posttraumatic stress disorder (PTSD)	A psychological injury characterized by reexperiencing trauma, hyperarousal, and withdrawal.
The Troubles	A term used to describe the 30 years of civil unrest and paramilitary campaigns, bordering on civil war, in Ireland.

Introduction

The *Oxford English Dictionary* defines terrorism as the use of violence and intimidation for political ends. This definition could apply to any kind of warfare. However, the term terrorism usually refers to unconventional tactics of warfare used by a less powerful group against a more powerful adversary. In this way, the Roman legions encountered terrorism in the woods of Germany. As a tactic, it has become more successful with the development of technology, so that a small number of terrorists with modern weapons can do a great deal more damage than many men wielding swords and halberds. The extent to which a small number of highly motivated men can inflict catastrophic pain on a mighty enemy was demonstrated by the events of September 11. Terrorism as a tactic is increasingly attractive to groups who lack political or military power.

One man's terrorist is another man's freedom fighter, and terms such as commando, guerrilla, partisan, and paramilitary help to confuse the picture further. At the same time, many former terrorists have been accepted into the political mainstream, and the progression from terrorist leader to statesman seems to be the rule rather than the exception (e.g., Nelson Mandela). Other terrorist organizations become increasingly involved in organized crime, and the pattern of terrorism is moving away from organized political groups to small, very extreme groups and even to individuals (e.g., the Uni bomber and the cell structure of the IRA and various Islamic fundamentalist groups). The field of terrorism, then, is a complex one, with the spectrum including a wide variety of separate phenomena.

Effects of Terrorism on the Terrorist

There has been little research from Northern Ireland examining the psychological profiles of active members of terrorist organizations. During the Provisional Irish Republican Army hunger strikes, hunger strikers were assessed and found not to be suffering from any psychiatric illness. Their self-harm was deemed to be a political act. Anecdotal information from professionals working in the prisons in Northern Ireland, suggests that various terrorist organizations are associated with different types of motivation and, hence, with different psychological profiles. For example, the members of some terrorist organizations are reputed to be very calm, calculating, politically motivated individuals, with very few of the traits that would normally be described as personality disorder. In contrast, the members of other organizations (particularly those organizations with an interest in organized crime) tend to have psychological profiles more typical of violent criminal offenders.

There is no evidence from Northern Ireland suggesting that those who engage in terrorist violence are more likely to be psychiatrically ill than members of the general population. Indeed, it seems that those who perpetrate terrorist violence, as a rule, tend to escape psychologically unscathed. Some individuals are troubled by their conscience, and the general public would like to believe that those who commit particularly horrific crimes will be troubled by these crimes for the rest of their lives. In general, however, it seems that 10 years after a horrific violent crime it is the victims and the bereaved who will wake up screaming in the night, whereas the perpetrator sleeps soundly.

There are positive psychological aspects that help to motivate terrorists. The terrorist act can act as a form of catharsis, relieving feelings of anger and frustration that have often built up over many years. At the same time, terrorists usually see themselves as martyrs. This is a constant theme, from the blood

sacrifice of Patrick Pearse in 1916, to the ideas of martyrdom among Islamic suicide bombers today.

These positive aspects are, for the terrorist, balanced by the threat that he or she poses to family and friends, the fear of living in secret, and the ultimate fear of losing his or her life. Thus, a young man who is considering becoming a terrorist is contemplating a tactic that is both highly effective and highly costly.

Effects of Terrorism on Victims

The effects of terrorist violence on victims has been described in detail (see **Northern Ireland, Post Traumatic Stress Disorder in**). Research performed with survivors of violence in Northern Ireland indicates that normal individuals subjected to extreme terrorist violence are likely to develop posttraumatic stress disorder (PTSD) and that normal individuals who are bereaved by terrorism are likely to develop a grief reaction. We use arbitrarily defined time points to describe the syndromes of delayed and chronic PTSD and to define delayed and chronic grief reactions. However, all of these psychological injuries, which happen to normal people put under intolerable strain, fall within a continuous spectrum, ranging from a psychological upset lasting a few days to symptoms that can be lifelong.

Vulnerable individuals may also be the victims of terrorist violence. Under these circumstances, the violence can act as a nonspecific stressor or life event precipitating an episode of psychiatric illness. For example, individuals prone to depression or anxiety may develop an anxiety neurosis or depressive neurosis if exposed to violence. Clearly terrorist violence has a devastating effect on the lives of those exposed to it, whether they are psychologically vulnerable or not, resulting in incapacitating physical and psychological injuries, as described by the work of Curran and colleagues, among others, working in Northern Ireland.

A high percentage of individuals who experience serious violence develop PTSD. Studies across the world, looking at different types of traumatic incidents, indicate that between one-quarter and one-half of individuals who experience serious trauma develop PTSD. It seems that the key factor in developing this condition is the fact of being made aware of one's own mortality at a very deep level. For this reason, it seems that it is the individual's perception of the event that causes the psychological injury, and because of this, different individuals can respond differently to the same type of trauma. For example, if two friends are caught up in a terrorist bomb blast, one might develop PTSD and the other might escape relatively unscathed. For the same reason, similar traumatic events may affect the same individual differently at different times. For example, a bus driver may be hi-jacked at gunpoint on four occasions without developing any psychological problems, only to develop severe PTSD when hi-jacked for a fifth time under virtually identical circumstances.

Anyone who has been exposed to serious violence may develop the core symptoms of PTSD, namely reliving the incident, hyperarousal, and emotional constriction. These core symptoms are seen after a wide variety of traumatic situations in different parts of the world. In addition to the core symptoms, an individual may develop a number of additional symptoms under specific circumstances. In situations of terrorist violence, hostage taking and torture are unfortunately frequent occurrences. This type of prolonged trauma tends to result not only in PTSD but can also in a syndrome similar to learned helplessness. Individuals recovering from traumatic events of this kind often experience profound symptoms of depression. Survivors of hostage taking often have these symptoms mixed with feelings of relief to be alive and a paradoxical gratitude to their captors for allowing them to live.

One common form of terrorist violence in Northern Ireland has been assassination attacks on the security forces launched from private houses. In these attacks, families are often held at gunpoint overnight while an ambush is set up. Families who experience these attacks often develop symptoms of PTSD associated with symptoms of depression and learned helplessness. The situation is complicated further by a group of symptoms that are often seen after burglaries. These symptoms include obsessional cleanliness and a feeling that the home has been violated, making it difficult to relax in the house where these events occurred.

Free-floating anger and irritability are core symptoms of PTSD. After natural disasters or tragedies for which no one is to blame, this free-floating anger often takes the form of irritability, and the survivors of natural disasters often direct their anger and hostility toward those closest to them. In the case of terrorist violence, there is a natural focus for this free-floating anger and victims understandably direct their anger toward the terrorist organization. In some circumstances, this can be helpful because there is less tendency for anger to be directed inward or toward other family members.

Individuals who have suffered terrorist violence in their own homes find it very difficult to recover from their psychological injuries while they continue to live in the environment where the trauma occurred. Every time they walk into a certain room they are reminded

of the violence that occurred there. At the same time, survivors of this type of violence often feel so much anger toward the perpetrators of the violence that they are determined not to be forced out of their homes. They refuse to run away. This appears to be a decision of principle, but it can tend to delay the natural healing process. Security lights and alarms are a nightly reminder of the violence that took place in the home. Those individuals who swallow their pride and move to a new environment often recover more quickly from their psychological injuries.

Even after the paramilitary ceasefires in Northern Ireland, there have been a number of terrorist incidents, most notably a bombing outrage in the town of Omagh in which large numbers of civilians were killed. This incident, in particular, generated a number of research studies on PTSD and specifically on the treatment of PTSD. As a result, the Department of Health in Northern Ireland produced the *CREST Guidelines for the Management of Post Traumatic Stress Disorder in Adults* in June 2003. This advice highlights the importance of cognitive-behavioral therapy (CBT) and eye movement desensitization, along with symptomatic treatment with antidepressants.

The Effect of Terrorism on Society

The effects of terrorist violence on a society are rather more difficult to define. As mentioned earlier, terrorism can take many forms, from the organization of mass protests and violence to individual attacks on selected targets. Curran and colleagues have pointed out that the form of civil violence over the years in Northern Ireland changed from mass disturbances, affecting large numbers of people, to more surgical attacks on selected targets, affecting only a very small number of victims. Community studies, examining indicators of mental health, have failed to demonstrate a greater psychiatric morbidity among the population of Northern Ireland. This has been explored in some detail by Cairns and Wilson, who examined various indicators of psychiatric morbidity, including psychiatric admission and referral rates, suicide and attempted suicide, substance abuse, clinical studies, and community studies. The overall impression is that during The Troubles it was remarkably difficult to demonstrate a higher level of psychiatric morbidity among the general population of Northern Ireland. Loughrey and coworkers postulated that this may have been because of the changing nature of violence in Northern Ireland to more selective targeting. The victims of violence, although affected severely, were perhaps too small in number to be detected by community surveys.

In the early years of The Troubles, Lyons described a year by year decrease in the suicide rate in Northern Ireland with the outbreak of civil violence. He believed that this might be explained by Durkheim's theory of anomie. In the late 1800s, Durkheim predicted that, during periods of war, suicide rates would decrease because the aggression of the population was being externalized toward a common enemy. This prevented aggression from being directed inward in the form of self-harm. Lyons's original results seemed to support this. However, when the suicide rates in Northern Ireland were examined over a more extended period of time by Curran and colleagues, it was found that the increase in the suicide rate at the start of The Troubles, observed by Lyons, may have simply been part of a wide year-to-year random variation. Durkheim's theories pertaining to social cohesion may still be relevant, however. Curran and coworkers postulated that the detrimental effect on the mental health of a small number of victims directly affected by terrorist violence might be balanced by a beneficial effect of polarizing sectarian community spirit on a larger number of marginalized individuals within society at large. In other words, although the direct victims of violence suffer greatly, they are small in number in most urban terrorist campaigns. At the same time, polarized sectarian communities, such as those in Northern Ireland, though deplorable in many ways, may paradoxically offer increased social support to marginalized individuals from within their own community. It is difficult to test this hypothesis, but most psychiatrists in Northern Ireland will be aware of some patients whose level of psychological functioning seems to be improved because of the increased social cohesion of the frontline communities in which they live.

Presumably the function of terrorism is to evoke terror. Initially, terrorists in Northern Ireland seemed to have been singularly unsuccessful in this regard. The terrorist campaign went on so long that the population of Northern Ireland viewed the risk of terrorist violence as just another of the statistical risks that face us when we get out of bed each morning. It almost seems that there is a homeostatic mechanism in the human psyche with regard to risk, such that exposure to risk develops tolerance. This seems to apply to risk of all sorts. The population of the Western world has become surprisingly blasé about the risks associated with motor vehicles, despite the huge carnage produced by these machines on our roads. Similarly, most individuals within society accept the risk of violent crime when they leave their homes. These mechanisms for living with risks appear to be largely based on denial and associated with thoughts such as "it will never happen to me."

Some groups within Northern Ireland, such as members of the security forces who are particularly at risk of terrorist violence, develop a gung-ho culture, where risk is denied completely. This is often associated with substance abuse in order to maintain the high levels of denial. There is a saying in Northern Ireland: "You can get used to anything." In the same way that money does not make you happy, it seems that terror does not necessarily make you afraid.

A number of studies within Northern Ireland have demonstrated a rebound of psychiatric morbidity after the declaration of the various terrorist ceasefires. However, this increase in psychiatric morbidity appears to be relatively short-lived. It may be that the atmosphere of peace allows people to come forward and seek help for psychological problems that they had chosen to ignore while the violence was continuing. The reaction of the community to terrorist violence can be seen as mirroring an individual's reaction to trauma. In traumatic circumstances, many individuals continue to function in a calm, detached manner while the violence is occurring, only to find themselves breaking down some hours later after they have had time to think about the dangers of the situation. This seems to be a natural adaptive mechanism. When survival is the first priority, individuals get on with what needs to be done at the time. They postpone thinking about the horror of the situation until it is safe to do so. It seems that societies function in a similar way, using denial mechanisms in order to continue functioning. However, the reality of the situation comes flooding in when peace is restored, resulting in the rebound of psychiatric morbidity.

Terrorism is here to stay. Just as individuals become inured to violence and tend to return to the state of normal denial, so societies learn to live with the chronic risk of terrorist violence and it becomes just one of the many risks in a modern world. When this happens, terrorism no longer succeeds in evoking terror. It becomes a less powerful tactic, and it becomes less likely that young men and women will be prepared to pay the terrible price of becoming terrorists.

In summary, the effects of terrorist violence are not as we would imagine them to be in a fair world. Most active participants in violence are psychologically unscathed. The lives of victims are often devastated, whereas the level of psychological morbidity in a community exposed to terrorism is lower than might be expected. In the end, the damage done by terrorism to a community appears to be spiritual rather than psychological.

See Also the Following Articles

Northern Ireland, Post Traumatic Stress Disorder in; Posttraumatic Stress Disorder – Clinical; War, Suicide and Sacrifice.

Further Reading

Bell, P., Kee, M., Loughrey, G. L., et al. (1988). Post traumatic stress disorder in Northern Ireland. *Acta Pyschiatrica Scandinavia* **77**, 166–170.

Cairns, E. and Wilson, R. (1984). The impact of political violence on mild psychiatric morbidity in Northern Ireland. *British Journal of Psychiatry* **145**, 631–635.

Clinical Research Efficiency Support Team (2003). *CREST guidelines on the management of post traumatic stress disorder in adults*. Belfast, UK: Department of Health and Social Services.

Curran, P. S., Finlay, R. J. and McGarry, P. J. (1988). Trends in suicide in Northern Ireland 1960–86. *Irish Journal of Psychological Medicine* **5**, 98–102.

Loughrey, G. L., Curran, P. S. and Bell, P. (1993). Post traumatic stress disorder and civil violence in Northern Ireland. In: Wilson, J. P. & Raphael, B. (eds.) *The international handbook of traumatic stress syndrome*, pp. 377–380. New York: Plenum.

Lyons, H. A. (1972). Depressive illness and aggression in Belfast. *British Medical Journal* **1**, 342–344.

Terrorism: Conseqences of *See:* Crime Victims; Impact of Terrorism on the Development of Mental Health Symptoms; Lockerbie Air Crash, Stress Effects of; Suicide Terrorism, Genesis of; Terrorism; War, Suicide and Sacrifice.

Terrorism: World *See:* Firefighters, Stress in; Religion and Stress.

Testosterone *See:* Androgen Action.

Thermal Stress

C M Blatteis
University of Tennessee Health Science Center, Memphis, TN, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by C M Blatteis, volume 3, pp 581–584, © 2000, Elsevier Inc.

Thermal Environments
Physiological Mechanisms of Heat Exchange
Metabolic Heat Production
Response Patterns to Sustained Thermal Stress
Other Environmental Factors Influencing Thermal Stress
Pathophysiological Consequences of Thermal Stress
Thermal Stress Not Due to Exposure

Glossary

Acclimatization Autonomic or behavioral changes occurring within an organism that reduce the strain caused by stressful changes in the natural climate. Acclimation refers to the adaptive changes that occur within an organism in response to experimentally induced changes in a particular climatic factor (e.g., ambient temperature) in an artificial, controlled environment.

Autonomic temperature regulation The regulation of body temperature by involuntary thermoeffector responses to heat and cold that modify the rates of heat production and heat loss. Autonomic in this sense does not mean that all responses are controlled by the autonomic nervous system.

Body heat balance The steady-state relation in which the heat produced by the body equals that lost to the environment.

Brown adipose tissue (BAT) A particular form of adipose tissue with selective distribution that differs both structurally and functionally from white adipose tissue. Its main role appears to be regulatory thermogenesis. Its activity is controlled by the sympathetic nervous system and various humoral factors, especially catecholamines.

Core temperature The mean temperature of the thermal core (i.e., the interior of the body), usually represented by the temperature of a particular internal site, for example, rectal or oral.

Thermal strain Any deviation of body temperature induced by sustained thermal stress that cannot be fully compensated by temperature regulation. Also, any activation of thermoeffector activities in response to thermal stress that causes sustained changes in the state of other, nonthermal regulatory systems.

Thermoeffector An organ system and its function that affect heat balance in a controlled manner as part of the processes of temperature regulation.

Thermal Environments

Thermal stress is defined as any change in the thermal relation between a temperature regulator and its environment that, if not compensated by thermoregulatory effector reactions, will result in hyper- or hypothermia. A temperature regulator, thus, is an organism that has the capacity to appropriately activate coordinated, autonomic, and/or behavioral heat conservation or loss (thermolytic) and heat production (thermogenic) effectors such that its body (core) temperature (T_c) remains relatively stable even in thermal environments that, *a priori*, might be expected to cause T_c to rise above (hyperthermia) or fall below (hypothermia) its normal range. Animals that possess this capability are called homeotherms. The condition in which T_c is within ± 1 SD of its normal range is termed cenothermy. The thermal environment in which neither the mechanisms for extra heat production (e.g., shivering) nor for extra heat loss (e.g., sweating) are activated is termed the thermoneutral zone; under these conditions, the T_c of healthy subjects at rest does not vary more than $\pm 2^{\circ}$C around its 24-h mean. Acute exposure to an environment warmer than thermoneutrality results in more heat being gained from the environment than lost to it; conversely, more heat is lost than produced by the body on exposure to an environment cooler than thermoneutrality. From a thermal physiological point of view, inactive humans are tropical animals when naked, thermoneutrality being an ambient temperature (T_a) of 28–30°C. Thermal stress, therefore, begins as soon as T_a is higher or lower than this range. However, homeotherms may be heat-stressed at low T_a if they are active.

Physiological Mechanisms of Heat Exchange

The unavoidable consequence of heat (or cold) exposure is that unless more heat were lost from the body

(or conserved and produced by it), the body-heat balance could not be maintained and, hence, an excessive rise (or fall) of T_c could potentially occur. However, the relative stability of T_c under most thermal conditions indicates that both heat production and heat loss are regulated and that, furthermore, temperature sensors exist throughout the body that measure its temperature continuously. In the skin, the detection of warm and cold sensations has recently been ascribed to the activation of specific, temperature-activated transient receptor potential ion channels (thermoTRPs) located within the free endings of cutaneous sensory nerves. The nature of the internal body thermosensors are not yet clarified. This information is transmitted to the brain, where it is integrated and transduced into signals to thermoeffectors that are then activated so that appropriate countermeasures are initiated. These responses involve both behavioral and autonomic processes. The former modulate primarily the effective heat-exchanging surface above the skin and are expressed by volitionally or instinctively changing the skin's microclimate (e.g., by postural changes, moving to a different thermal environment, or wetting the skin). Behavioral adjustments generally precede autonomic responses; they are more economical energetically and, often, sufficient. Autonomic adjustments involve factors below the skin. They control the patency of the small arterioles that deliver blood (i.e., heat) to the surface by means of sympathetic vasoconstrictor and vasodilator systems and various cotransmitters. When open (cutaneous vasodilation), the warm arterial blood passes through these, warming the skin. As long as the skin temperature remains higher than the ambient, heat is dissipated into the environment. Conversely, when these channels are closed (cutaneous vasoconstriction), the bloodstream is conveyed through deeper arteriovenous anastomoses such that heat does not reach the surface so readily, reducing the skin-to-ambient temperature gradient and conserving body heat. This is aided further by the fact that the deep, larger vessels of the limbs lie in parallel so that warm arterial blood flowing toward the periphery passes close to cooler, returning venous blood flowing in the opposite direction, thus allowing heat to be transferred by countercurrent heat exchange from the arterial to the venous flow. This precools the arterial blood, thereby contributing to increasing the thermal insulation of the extremities.

When ambient temperature exceeds skin temperature, the body gains heat. Under these conditions, the only means by which heat can be dissipated is by evaporation. For this purpose, animals that possess eccrine sweat glands activate them; those that do not, pant. The efficacy of this process is dependent directly on the ambient and skin vapor pressure difference; that is, when the ambient relative humidity is high, little evaporation occurs. Sweating is much more expensive metabolically than cutaneous vasodilation. It also results in large losses of water and sodium, which are potentially dangerous.

Metabolic Heat Production

The heat released when the organism is resting at thermoneutrality is called obligatory nonshivering thermogenesis (NST-O); it is derived from the metabolic activity of the body. T_c is maintained within its narrow range by the controlled loss of heat from this activity. The skeletal muscle mass makes up the greatest mass of the body (~45%), but at rest it produces only 18% of the NST-O. However, more energy and, hence, heat are expended when the muscles actively contract, in direct proportion to the external work done (e.g., as in exercise). However, if no external work is performed by the contracting muscles, most of the energy expended is given up as heat. This is, in fact, what occurs in a cold environment by means of shivering. However, shivering is fatiguing, not terribly efficient (~11%), and cannot be sustained indefinitely. Another thermogenic process that does not involve muscular contractions, namely thermoregulatory nonshivering thermogenesis (NST-T), is activated in such cases. The principal effector organ of NST-T is BAT. Its thermogenic capacity is derived from the collapse of the electrochemical proton gradient normally generated by cellular respiration, limiting adenosine triphosphate synthesis, the energy currency of the body. Consequently, the energy expended is given up as heat and conducted to the numerous vessels permeating this tissue and is convected by the bloodstream throughout the body. A specialized protein, uncoupling protein 1 (UCP1, or thermogenin), present in the mitochondrial inner membrane of BAT adipocytes, facilitates this process when activated by the sympathetic nervous system. BAT NST-T is particularly prominent in neonates, but recedes with age and size. It develops over time in cold-exposed adult rodents and other small animals. In neonates and small animals, NST-T and shivering are meshed reciprocally.

Response Patterns to Sustained Thermal Stress

If a person remains exposed, unprotected, for a prolonged time to the natural climate, he or she would be at risk of eventually becoming physically exhausted from the continuous shivering in the cold or dehydrated and hyponatremic from the extensive sweating

in the heat. To prevent such noxious results, the body institutes thermoadaptive modifications that reduce the physiological strain caused by prolonged stressful thermal exposures. These adaptations, termed acclimatization, occur over an extended time course, measured in days and longer. The gradual replacement of shivering by NST-T in long-term cold exposure of small animals has already been mentioned. This type of response is called metabolic adaptation. Some homeotherms react to prolonged cold by entering into a state of lethargy called hibernation. Adult humans acclimatize to cold by using appropriate behaviors (clothing, heaters, etc.). The autonomic analog in most animals is the growth of thicker fur and the deposition of subcutaneous fat (insulative adaptation). Inadequately insulated, chronically cold-exposed humans may also develop a lower T_c and a lesser fall in skin temperature (i.e., a warmer skin; hypothermic adaptation) in time.

In long-term heat exposure, the sweat glands develop the capacity to secrete more water and sooner (i.e., they become activated at a lower threshold T_c over time); the amount of salt lost in sweat from the body also is reduced. These effects can be achieved by 1–4 h of heat exposure for 5–7 days and are mediated by the increased secretion of antidiuretic hormone and aldosterone; thirst is also stimulated. The process is facilitated by physical exercise, which, by itself, enhances the sweat rate, and persists as long as the heat exposure continues; afterward, it gradually wanes. Analogous to hibernation in the winter, some homeotherms enter into a state of summer lethargy called estivation. During moderate chronic heat exposure, the onset of both sweating and skin vasodilation occur at a higher T_c (hyperthermic adaptation).

Systemic adaptation to thermal stress also extends to the cellular level (thermotolerance). Thus, cells within the body core (e.g., exercising skeletal muscle) or in its shell (e.g., skin cells) that are acutely exposed to intense, but not lethal, heat or cold react to these stressful stimuli by producing specialized proteins (heat shock and cold shock proteins) that serve to protect them against a subsequent, otherwise potentially lethal heat or cold exposure. In contrast to heat or cold acclimatization, however, thermotolerance develops rapidly (within hours) but is temporary (3–5 days).

Other Environmental Factors Influencing Thermal Stress

In nature, a variety of factors are present that aggravate the risk of thermal stress occurring sooner than due to a T_a change alone. These include water, due to its high thermal conductivity (beneficial in the heat, detrimental in the cold, as in wet-cold, e.g., chilblains and immersion foot); blowing wind (windchill factor), by thinning the thickness of the thermal boundary layer, a thin film of warmed air that surrounds the body (beneficial in the heat, detrimental in the cold); and barometric pressure, by affecting heat exchange by convection (reduced at high altitude) and evaporation (enhanced at high altitude) and by affecting the partial pressure of atmospheric gases (e.g., the metabolic response to cold is reduced at altitude). Other environmental factors, such as the relative humidity and the intensity of solar radiation, also contribute to the imposition of thermal stress.

Pathophysiological Consequences of Thermal Stress

During periods of heavy sweating, a strain is imposed on the heat-exposed individual's homeostatic systems that, if not properly compensated, can cause the development of symptoms (heat illnesses). Similarly, when the various defense reactions are inadequate to cope with the stress of cold exposure, severe local tissue injury (cold injuries) may develop. The risks of hypo- and hyperthermia were mentioned earlier. The susceptibility to thermal stress increases with age.

Thermal Stress Not Due to Exposure

Thermal stress may also occur due to factors not involving exposure. For example, as already mentioned, exercise can induce heat stress by increasing the rate of heat gained (i.e., produced) in excess of that lost. Conditions of decreased or increased T_c may be induced, even in thermoneutrality, by drugs that impair normal thermoregulatory functions (iatrogenic hypo- or hyperthermia), by idiosyncratic drug reactions (e.g., malignant hyperthermia), or secondary to underlying illnesses that compromise the body's thermoregulatory mechanisms (e.g., peripheral vascular disease, paralysis, or endocrinopathies). Fever, however, is not a thermal stress because it represents a regulated, deliberate T_c rise resulting from the active operation of thermogenic effectors and not the unavoidable consequence of the passive gain of heat in excess of the capability of active thermolytic effectors to dissipate it; it is a defensive response to the invasion of the body by infectious pathogens. A regulated fall in T_c also occurs, particularly in rodents exposed to toxic chemicals and hypoxia/ischemia; it protects the afflicted host from the consequent pathological effects of these insults.

See Also the Following Articles

Environmental Factors; Environmental Stress, Effects on Human Performance; Febrile Response; Heat Resistance; Heat Shock Response, Overview; Hyperthermia; Hypothermia; Temperature Effects; Thermotolerance, Thermoresistance, and Thermosensitivity.

Further Reading

Blatteis, C. M. (ed.) (1998). *Physiology and pathophysiology of temperature regulation*. Singapore: World Scientific.

Bligh, J. (1998). Mammalian homeothermy: an integrative thesis. *Journal of Thermal Biology* **23**, 143–258.

Cannon, B. and Nedergaard, J. (2004). Brown adipose tissue: function and physiological significance. *Physiological Reviews* **84**, 277–359.

Charkoudian, N. (2003). Skin blood flow in adult human thermoregulation: how it works, when it does not, and why. *Mayo Clinic Proceedings* **78**, 603–612.

Fregly, M. J. and Blatteis, C. M. (eds.) (1996). *American Physiological Society handbook of physiology. Sec. 4: Environmental physiology* (2 vols.). New York: Oxford Univ. Press.

Glossary of terms for thermal physiology (2001). *Japanese Journal of Physiology* **51**, 245–280.

Kenney, W. L. and Munce, T. A. (2003). Invited review: aging and human temperature regulation. *Journal of Applied Physiology* **95**, 2598–2603.

Kregel, K. C. (2002). Heat shock proteins: modifying factors in physiological stress responses and acquired thermotolerance. *Journal of Applied Physiology* **92**, 2177–2186.

Patapoutian, A., Peier, A. M., Story, G. M., et al. (2003). ThermoTRP channels and beyond: mechanisms of temperature sensation. *Nature Reviews* **4**, 529–539.

Pozos, R. S. (ed.) (2001). Sec. II: Cold environments, In: Pandolf, K. B. & Burr, R. E. (eds.) (2001). *Textbooks of military medicine: medical aspects of harsh environments* Washington, DC: U.S. Department of the Army, Office of the Surgeon General.

Reid, G. (2005). ThermoTRP channels and cold sensing: what are they really up to? *Pflugers Archiv – European Journal of Physiology* **451**, 250–263.

Sell, H., Deshaies, Y. and Richard, D. (2004). The brown adipocyte: update on its metabolic role. *International Journal of Biochemistry and Cell Biology* **36**, 2098–2104.

Stocks, J. M., Taylor, N. A., Tipton, M. J., et al. (2004). Human physiological responses to cold exposure. *Aviation, Space and Environmental Medicine* **75**, 444–457.

Wenger, C. B. (ed.) (2001). Sec. I: Hot environments, In: Pandolf, K. B. & Burr, R. E. (eds.) (2001). *Textbooks of military medicine: medical aspects of harsh environments* Washington, DC: U.S. Department of the Army, Office of the Surgeon General.

Thermoregulation *See:* Febrile Response; Thermotolerance, Thermoresistance, and Thermosensitivity.

Thermotolerance, Thermoresistance, and Thermosensitivity

S Gatti McArthur, D Alberati and T Bartfai
Hoffmann-LaRoche, Basel, Switzerland

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by S Gatti McArthur, D Alberati, T Bartfai, volume 3, pp 585–594, © 2000, Elsevier Inc.

Temperature Control
Thermoresistance
Thermotolerance
Thermosensitivity
Concluding Remarks

Glossary

Adaptation/ habituation	A modification of regulatory processes taking place during prolonged or repeated exposure to extreme ambient temperatures.
Poikylothermy (PKT)	The typical metabolic condition of ectothermic animals that use external sources of heat to increase body temperature. In endotherms, a status similar to PKT can be observed in extreme ambient conditions or under pharmacological treatment with ethanol or volatile anesthetics.

Thermoneutrality (TN) A range of ambient temperatures. Within the TN range, it is not necessary for the organism to activate warming or cooling mechanisms to keep the core body temperature constant. The TN range is different in different species and strains. Furthermore, there is a high degree of interindividual variability, and these differences could be related to differences in diet, age, and body weight. A TN range can be determined experimentally by measuring O_2 consumption as an indirect index of the rate of oxidative metabolism. More precise analysis requires biocalorimetry measurements to determine heat production (HP) and heat exchange (HE) in the different regions/tissues/cells of the body. Because rapid eye movement (REM) sleep is also sensitive to the ambient temperature, the TN range can be determined by measuring the total REM sleep time (longer in the TN zone).

Thermopreferendum The temperature (T) range selected by microorganisms or animals placed in a thermocline. The association of a certain T range with the presence of food or memory of previous ambient conditions can influence the selection by the organism. Behavioral tests based on the resetting of the thermopreferendum can be used to demonstrate T set point shifts.

Thermoresistance Hormonal, metabolic, and cardiovascular reactions to environmental temperature changes that allow the organism to keep a constant control on core body temperature in the presence of unfavorable ambient conditions.

Thermosensitivity Capacity to respond to temperature changes. Thermosensitive (Ts) neuronal units modify the number of pulses per minute in response to small changes in temperature. Based on the response to direct warming or cooling procedures, warm-sensitive (pulses/10 min = $Q_{10} > 1.2$) or cold-sensitive ($Q_{10} < -0.6$) neurons can be distinguished, respectively.

Thermotolerance Mechanisms of adaptation that allow survival in extreme ambient conditions, for example, short-term changes in heat production and vascular control of heat exchange and long-term changes in membrane composition and insulation capacity.

Uncoupling effect When the protonic gradient across the inner membrane of mitochondria is partially or totally dissipated and heat production (HP) takes preference over ATP production. Uncoupling proteins and uncoupling agents (e.g., dinitrophenol) cause intracellular HP. In the presence of uncouplers, HP is proportional to O_2 consumption, percentage of uncoupling effect, mitochondrial membrane potential, and duration of the effect.

This article briefly reviews the mechanisms controlling heat production (HP) and heat exchange (HE) in endotherms (in mammals in particular) able to adapt to different, and sometimes extreme, environmental conditions. Hibernation (and estivation) represents a classic example of extreme control on thermoregulation. When food is scarce, lowering the core body temperature (CBT) allows the survival of small mammals and reptiles with minimal energetic costs. Introductory concepts on temperature (T) control in endotherms are also given.

Temperature Control

Temperature (T) control or, better, the control of the effects of T is a vital need for every cell type, even for prokaryotic organisms, given that protein conformation and membrane integrity are jeopardized at freezing, as well as at higher than 40°C temperatures (in aqueous environment that prevails *in vivo*). Exceptions to this rule in cryobiology require nonaqueous solutions. In bacteria, adaptive changes associated with small T variations (Δ: 2–10°C) always involve changes in cell membrane composition and involve the rapid induction of enzymes such as fatty acid desaturases. *Escherichia coli* responds to small changes in T by altering its growth rate, swimming behavior, and thermotaxis. Several membrane proteins (methyl-accepting chemotaxis proteins) are involved in the control of these events. The thermo/chemoreceptors are reversibly methylated at specific glutamate residues located in the C-terminal cytoplasmic domain and point mutations at the methylation sites can even invert bacterial thermotaxis. Nematodes (e.g., *Caenorhabditis elegans*) possess specific thermosensory neurons and interneurons organized in a network controlling thermotaxis. When placed in a thermal gradient, they display the ability to recognize and select the region where T is similar to the T at which they were cultured previously. Curiously, it is possible to distinguish between a cryophylic and a thermophylic behavior by selective laser ablation of one of the two interneurons controlling thermotaxis or by null mutations of either of two homeobox genes, *lin-11* or *ttx-3*, respectively. The neural circuit that regulates the response to changes in T in *C. elegans* has been characterized in detail. It consists of a thermosensory organ (neuron), two interneuronal populations, and motor neurons. Vertebrates

(amphibians and reptilians), such as lizards or frogs, exhibit a thermopreferendum when placed in a thermocline even if they are poikylothermic and thus unable to keep their core body temperature (CBT) constant in the same thermal gradient.

Only mammals and birds (warm-blooded animals) have developed a complex net of neurological, hormonal, cardiovascular, and metabolic systems that are able to keep a constant control on the CBT despite geographical, seasonal, or even diurnal differences in air T and water T. The possibility of keeping CBT constant represents an important step from an evolutionary point of view. At the cellular level, it implies that the rate of enzymatic reactions, membrane properties, and electrical excitability remain rather constant (and can be optimized) despite dramatic changes in heat production (HP) and heat exchange (HE) caused, for instance, by motor activity or cold exposure. In general, homeothermy may be one of the key evolutionary factors that have allowed uterine growth of fetuses and nocturnal motor activity.

Temperature Control in Homeotherms

Several homeostatic mechanisms allow the maintenance of a constant CBT in warm-blooded animals exposed to different ambient temperatures (AT): (1) the hormonal control on basal metabolic rate (BMR) (short term) and body mass (long term), (2) the hypothalamic/spinal neuronal network setting and controlling CBT, (3) the sympathetic and parasympathetic neuronal systems activated alternatively to govern effector mechanisms to warm or cool the body, and (4) the steady-state bulk of energy expenditure represented by the activities of muscular and adipose tissues.

Thermoresistance

Basal Metabolic Rate

When compared to poikylothermy (PKT) in animals of equal body mass, homeothermy requires the maintenance of higher BMR, with higher food intake and higher O_2 consumption. Indirectly, this means that homeotherms actuate behavioral and dietetic changes, adjusting food intake to seasonal T changes in order to guarantee sufficient energy storage. We should distinguish between (1) resting (basal) metabolic rate (MR; the heat produced is also called obligatory thermogenesis, T_g) and (2) inducible MR (the heat produced is also called facultative T_g). The resting MR can be largely modified by slow mechanisms of adaptation, whereas the inducible MR is activated quickly in order to react effectively to cold exposure. The major ways used to increase T_g rapidly are uncoupling of oxidative phosphorylation in mitochondria (nonshivering T_g) and muscular shivering. BMR in mammals is correlated positively to the ratio of body surface to body mass; animals with a higher surface/mass ratio require exponentially more energy to keep their CBT in the range of 36–37°C.

These dramatic differences in BMR should be considered when studying the adaptive mechanisms developed by different species to allow survival in extreme conditions. Thus, for example, the drastic reduction of CBT that accompanies hibernation is advantageous for small animals (e.g., squirrels) because they save much more energy proportionally than larger animals, such as bears. Passive mechanisms of thermotolerance, such as the reduction of HE because of increased insulation, do not modify the BMR. They simply expand the range of AT compatible with survival. Food digestion is another source of T_g, and different species change their diets in relation to the amount of heat required in different seasons.

Hormonal Control on Basal Metabolic Rate

The thyroid hormones, thyroxine (T_4) and triiodothyronine (T_3), play a central role in controlling cellular ATP levels, MR, and T_g during both fetal and adult life. T_3 and T_4 increase obligatory T_g and work in synergy with catecholamines to trigger facultative T_g. Furthermore, thyroid hormones are involved in embryogenesis, particularly in the differentiation of brown fat tissue (BAT). Studies have clarified the molecular mechanisms sustaining the different metabolic effects and the modulatory actions of thyroid hormones on insulin, glucagon, growth hormone, and adrenaline effects. Thyroid hormones enhance or suppress gene transcription via the specific interaction with different nuclear receptors. The effects of T_3 and T_4 on T_g are related to the induction of several classes of proteins that include uncoupling proteins (UCPs); FAD-linked glycero-3-phosphate dehydrogenase, a carrier transferring electrons from NADH to ubiquinone in complex I in the respiratory chain of mitochondria; and adenine nucleotide translocase, the ATP/ADP exchanger, which is also present in the inner mitochondrial membrane, where it controls the concentration of phosphate acceptor (ADP).

Table 1 shows genes whose expression is known to be under the control of thyroid hormones. UCP-1 protein expression is induced in rat BAT by cold exposure and suppressed by fasting; and T_3 increases UCP-1 production both by increasing gene transcription and by stabilizing mature UCP-1 mRNA. In BAT, local control of the relative levels of T_3 and T_4 is maintained via the presence of thyroxine 5′-deiodinase. This cold-inducible enzyme converts T_4 to the

Table 1 Genes involved in induction of thermotolerance or thermoresistance under the transcriptional control of thyroid hormones[a]

Mitochondrial proteins	*FAD-dependent glycerol-3-phosphate dehydrogenase (ATP/ADP exchange activity)*
	Tom20: outer membrane receptor
	Matrix heat shock protein 70
	Adenine nucleotide translocase (ATP/ADP exchange activity)
	Acyl-CoA oxidase
	Na^+K^+-ATPase (α and β subunits)
	Acetyl-CoA carboxylase
	Uncoupling proteins (UCP-1,-2,-3)
	ATP-dependent protease
Hormones	Renin
	Steroidogenic acute regulatory protein
	Oxytocin
	Prepro-TRH (negative control)
Other enzymes	Cholesterol-7α-hydroxylase
	Malate dehydrogenase (negative control)
	Cytochrome P450 (negative control on CYP4A1, CYP4A3, and CYPA23)
	Ca^{2+}-ATPases (in sarcoplasmic reticulum: SERCA1 and SERCA2)
	Phosphoenolpyruvate carboxykinase
Membrane receptors and ion channels	Glucagon receptor
	β1 and β3 adrenoreceptors
	Potassium channels: Kv1.2, 1.4, 1.5, 2.1, and 4.2 α subunits
Nuclear factors	Steroid receptor coactivator-1
	Cold-inducible coactivator of nuclear receptors
Others	Glucose transporter GLUT4
	Myosin isoforms
	Phospholamban: a major target of the β-adrenergic cascade in the heart

[a] SERCA, sarcoplasmic reticulum Ca^{2+} Atpase; TRH, thyrotropin-releasing hormone.

more potent T_3, thereby locally modulating the ratio between T_3 and T_4 levels. In muscles, thyroid hormones also control the expression of genes coding for myosin isoforms, the Na^+-K^+-ATPase pumps, and the Ca^{2+}-ATPase channels of the sarcoplasmic reticulum, sustaining, in this way, muscular performance during shivering.

Thyrotropin (TSH) production in the pituitary gland and, upstream, thyrotropin-releasing hormone (TRH) production in several hypothalamic and raphe nuclei control the synthesis and secretion of thyroid hormones. TRH also seems to exert additional actions in the suprachiasmatic nuclei (SCN), amygdala, lateral septal nuclei, and pineal gland. These actions are always related to chronobiological changes of the MR. During hibernation, the levels of thyroid hormones, TSH, TRH, and their receptors are decreased in all of the just-mentioned brain regions (except for TRH levels in the pineal gland).

In recent years, pharmacological therapies have been proposed to increase HP in obese patients, with special attention to the hypothalamic systems controlling the balance between energy storage and energy expenditure. Interesting observations were made in *ob*/*ob* and *db*/*db* obese mice, which lack the control of leptin (the starvation hormone) on adipose tissue metabolism. These animals exhibit a CBT that is 1–2°C below normal; several studies point to the almost total impairment of nonshivering T_g present in *ob*/*ob* torpid animals from the first days of life. This is probably due to (1) the defective regulation of thyroxine 5′-deiodinase, (2) the attenuated UCP mRNA expression in BAT during cold exposure, (3) the higher circulating levels of corticosterone observed in *ob*/*ob* mice, and (4) the increased sensitivity of these animals to leptin. Moreover, leptin can affect thermogenesis, modulating the sympathetic outflow directly as it activates neurons that coexpress cocaine- and amphetamine-regulated transcript (CART) and proopiomelancocortin (POMC) neurons that innervate sympathetic preganglionic neurons in the thoracic spinal cord. One of the hypothalamic nuclei constitutively expressing high levels of the leptin receptor long form is the arcuate nucleus. This nucleus contains NPY–glucose sensitive neurons, which represent a possible key link between leptin and T control because pro-NPY mRNA expression is inhibited by leptin and the direct injection of NPY into the anterior or dorsomedial hypothalamus causes deep hypothermia.

Hypothalamic and Brain Temperatures

Thermosensitive (Ts) neurons are present in the anterior brain stem of several endothermic vertebrates and even in ectotherms, further confirming a role for these neurons in the CBT control from a very

early stage in the phylogenetic scale. In mammals, the local warming or cooling of hypothalamic (anterior preoptic, among others) and spinal regions triggers panting, autonomic responses, and muscular shivering, respectively. Ts neurons are also activated by remote thermal stimuli. Approximately 25% of the neurons in anterior hypothalamus are Ts, and Ts neurons can also be found in various nuclei involved in thermal control (Figure 1), circadian rhythms (SCN), and sleep induction (diagonal band). The hypothalamic temperature (HT) plays a key role in the control of the CBT. The hypothesis of HT as an independent variable controlled locally and tightly, integrating peripheral and extrahypothalamic stimuli, is further sustained by the observation that T undergoes smaller fluctuations in hypothalamus

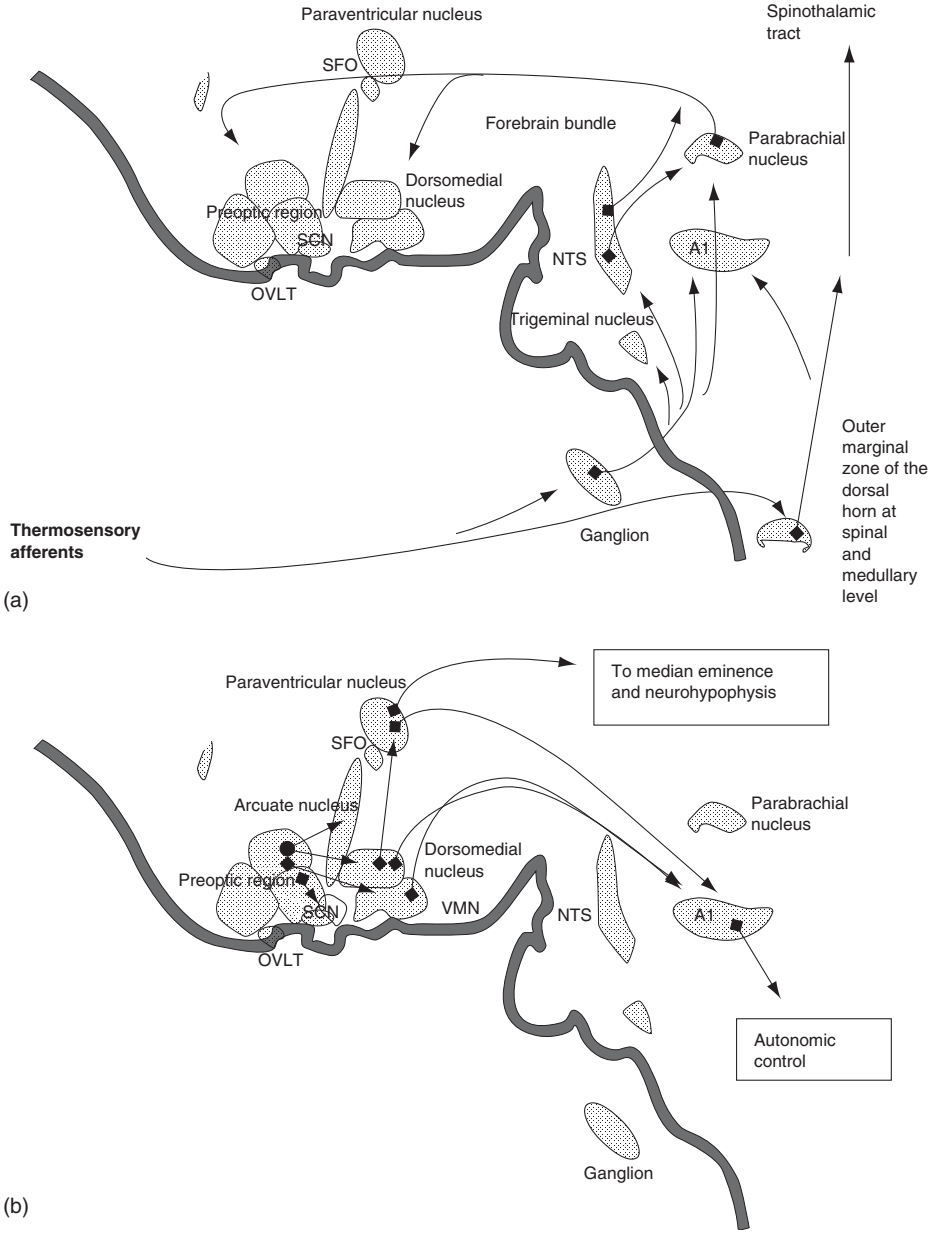


Figure 1 Schematic view of the principal thermosensory afferent and efferent systems in the rat brain. a, Afferent pathways; b, main descending efferent pathways. Hypothalamic regions and autonomic nuclei involved in core body temperature control are highlighted in a sagittal section of the hypothalamic/bulbar region. In a, periventricular hypothalamic organs – subfornical organ (SFO) and organum vasculosum lamina terminalis (OVLT) – are also indicated; their projections to the preoptic and paraventricular nuclei control the thermogenic response to circulating hormones. The suprachiasmatic nucleus (SCN), the principal circadian clock of the body, receives light information from the retina and projects to the autonomic division of the paraventricular nucleus. In b, the main hypothalamic descending efferent pathways control thermogenesis via the activation of the autonomic system. NTS, nucleus of the solitary tract; VMN, ventromedial nucleus.

than in other brain and trunk regions in response to several different stimuli. For instance, during hibernation, HT is set approximately 6–10°C higher than CBT. It is generally believed that local hypothalamic and brain heating is controlled by cardiovascular mechanisms, dependent on blood T and perfusion. Several studies were able to demonstrate that mammalian species can decrease the T of the brain below the T of the blood circulating in the large arteries of the trunk. This effect, namely selective brain cooling (SBC), is possible because of the presence of a carotid thin-walled rete that keeps the arterial blood in close contact with the cooler venous blood of the cavernous sinus. What is really under control is the T of the venous blood, which can be modified (sphincter of the angularis oculi vein), allowing the entrance into the sinus of venous blood coming from the upper respiratory tract and the facial vein. The activation of mechanisms of SBC is dependent on CBT and is usually evident when CBT is approximately 40–41°C. Again the onset and the degree of the effect are determined predominantly by T signals of hypothalamic origin. Although the carotid rete is not present in humans, several groups have suggested that SBC could also play a critical role in controlling brain T in humans, at least during physical activity and prolonged exposure to the sun. Endotherms, and mammals in particular, can count on a tight endocrine control of the HT set point, which can be modified in response to infection (fever), sleep, circadian and ultradian rhythms, pregnancy, and fertility cycles. The interaction among different stimuli is possible because of the multimodal response of Ts neurons to endogenous pyrogens, estrogens, glucose levels, osmolarity, and blood pressure, among other stimuli.

Autonomic System and Temperature Control

The major hypothalamic, bulbar, ganglionic, and spinal nuclei involved in the control of autonomic response during thermal stress are depicted in **Figure 1b**. Sympathetic activation can cause an immediate reduction of skin blood flow in thermoregulatory vessels, piloerection, heat production in BAT, muscular shivering, and increased respiratory rate. Parasympathetic control, however, is exerted mainly on exocrine sweat glands and cardiovascular and respiratory systems, respectively. The sympathetic system exerts a tight thermoregulatory control via the rhythmic and tonic activity of postganglionic neurons, whereas the integration of spinal and bulbar inputs is controlled by preganglionic neurons. Furthermore, every peripheral nerve contains silent units that can be activated in the presence of stress stimuli such as heat.

Adaptive changes in parasympathetic activity after physical exercise, especially in hot environments, cause bradycardia and increased exocrine sweat gland activity. In normal individuals, 2–4 million sweat glands are able to secrete approximately 3 l of hypotonic fluid/h. In athletes or acclimated people, the sweating rate is higher and starts as soon as there is heat load because of the lowered threshold of sweating response in hypertrophic glands.

Nonshivering Thermogenesis

In children and in many animal species, the major source of heat is BAT. The main role in heat production in BAT adipocytes is played by the mitochondrial protein thermogenin (UCP-1). UCP-1 may represent up to 15% of the total protein content of the inner mitochondrial membrane, and during 1 week of cold exposure the total content of UCP-1 in rat intrascapular BAT rises from 60 to approximately 300 μg, whereas the total protein content increases little. It has been calculated (using *in vitro* microcalorimetry) that 1 kg of stimulated BAT can produce approximately 300 W of heat energy. The amount of BAT, and of vessels supplying O_2 to the tissue, can be increased trophically during acclimation processes. Heat production in BAT is obtained by dissipation of the protonic gradient across the inner mitochondrial membrane. Studies indicate that the fatty acid shuttle of protons is responsible for this effect. The system is under the control of β_3 receptor activation in several species.

The general assumption of a lack of nonshivering T_g in adult humans has to be considered in light of the discovery of UCPs other than the BAT-UCP-1; UCP-2, UCP-3, UCP-4, and BMCP-1 are expressed differentially, mainly in the white adipose tissue, liver, heart, glycolytic muscles, and brain (UCP-4 and BMCP-1 are brain-specific). Structurally, these proteins are similar to other mitochondrial carriers, such as the oxoglutarate carrier. **Figure 2** shows a probable phylogenetic tree of this class of proteins based on the available cDNA sequences of UCPs from various species. Putative proteic sequences from *C. elegans* and yeast suggest the presence of these proteins also in lower evolutionary levels organisms. The analysis of the different UCP cDNAs shows that UCP1, UCP4, BMCP1, and plant UCPs possibly represent a more recent step from the evolutionary point of view, whereas UCP-2 may be phylogenetically older. The comparative analysis of the structure of the available UCP genes supports this hypothesis further in that the UCP2 gene does not contain a TATA box and is constitutively expressed and rather ubiquitous. UCP-2 and UCP-3 genes are contiguous, suggesting possible mechanisms of gene duplication. Tissue levels of the different UCP gene products are under translational control. Their expression is correlated with changes

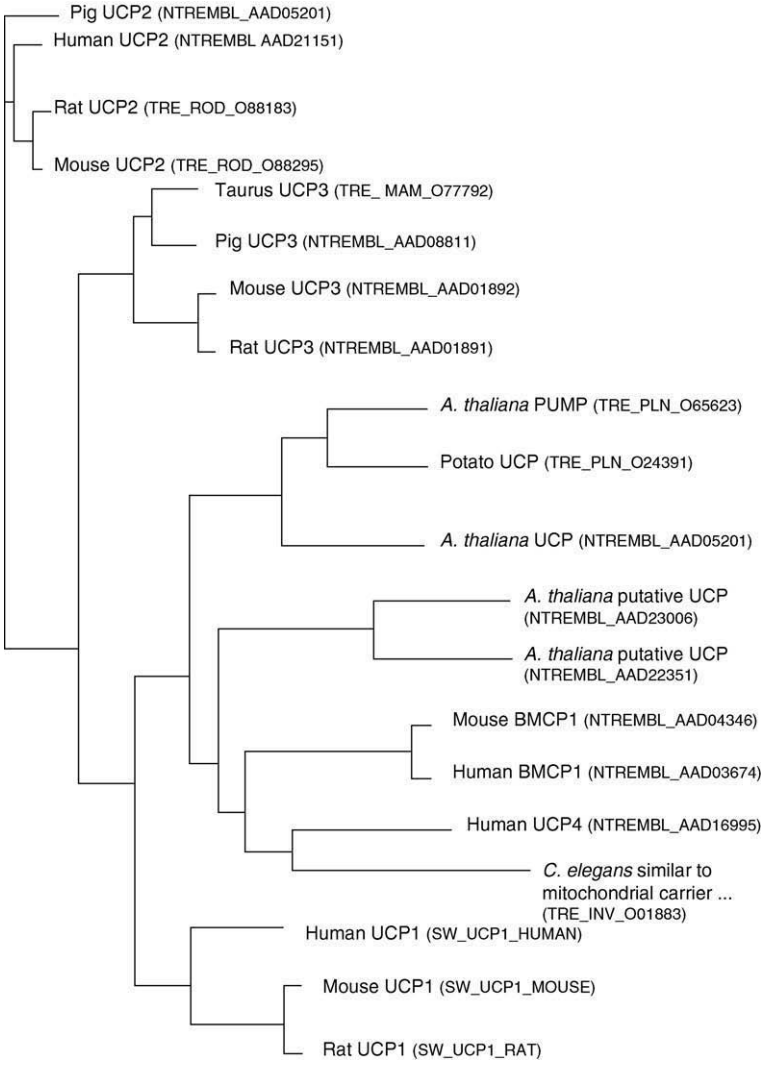


Figure 2 Phylogenetical tree of the uncoupling protein family(ies). The discovery of plant mitochondrial uncoupling protein (PUMP), of UCP-like proteins in *Arabidopsis thaliana* /potato, and of several novel mammalian uncoupling proteins suggests that the ancestral UCP gene has probably evolved prior to the divergence into animals and plants. This is also confirmed by the presence of a protein with uncoupling activity into the mitochondria of *Acanthamoeba castellanii,* a nonphotosynthetic protozoon, which in molecular phylogenesis appears on a branch basal to the divergence points of plants, animals, and fungi. The phylogenetic tree was obtained using the GCG software package.

in ATP production (lowered), inner mitochondrial membrane potential (lowered), and heat production (increased). However, their functional role in T_g is still under debate. UCP-2, UCP-4, and BMCP-1 mRNA expression in the brain and UCP-2 mRNA expression in the spleen, lymph nodes, and thymus suggest other possible physiological roles. Furthermore, UCPs are regulated differentially (possibly in a tissue-specific manner). The different affinity for free ATP/ADP and GTP/GDP nucleotides (inhibitors) could cause the activation of uncoupling at different metabolic states. Possible differences in terms of affinity to fatty acids and the influence of pH on nucleotide binding still require further characterization.

Muscular Shivering

Shivering is controlled by motor neurons and is caused by the repetitive contraction of both extensory and flexory muscles with a frequency directly proportional (on a logarithmic scale) to BMR. The Renshaw interneuron activation is responsible for the synchronization of motoneurons and for the upper discharge frequency. During cold exposure, more than one-third of the total heat produced can come from shivering, and muscular performance is dependent on the intracellular levels of several cold-inducible proteins, such as the carrier for glucose uptake into muscles (GLUT-4) and sarcoplasmic or endoplasmic reticulum Ca^{2+}-ATPases, respectively. Massive shivering

also allows a prompt reaction to intraoperative hypothermia or raises the CBT quickly to mount very high fevers (e.g., during renal infections).

Thermotolerance

Membranes

All cells possess the ability to modify the fluidity of the membranes in order to preserve structural and functional integrity in the presence of T changes. In poikylothermic animals, cold exposure increases the percentage of polyunsaturated fatty acids (mono- and polyunsaturated, mostly oleic acid) in phospholipids and the relative amount of phosphatidylethanolamine (PE) (vs. phosphatidylcholine, PC). These changes in the composition of the phospholipid bilayer preserve fluidity, probably because unsaturated fatty acids increase membrane disorder (homeoviscous adaptation). It should be noted that cold adaptation causes similar changes in the composition of the mitochondrial membrane in BAT with an increase in polyunsaturated fatty acids and PE and with a decrease in cholesterol content. The final effect is probably a greater membrane fluidity. The brain of hibernating animals contains less cholesterol and greater amounts of both PC and PE, monounsaturated fatty acids, arachidonic acid, and polar gangliosides.

Enzymes and Cryoprotective Polypeptides

The most impressive effect of hibernation is the drastic reduction of BMR. This effect is obtained by a tissue-specific inhibition of the activity of key metabolic enzymes such as pyruvate dehydrogenase. In some cases, a different phosphorylation state of the enzyme has been involved (6-phosphofructo-1-kinase); in other cases, the nonspecific effect of temperature change on the K_m for the different substrates seems to be sufficient to explain the observed effect. However, there is much more to learn in relation to cryophylic or thermophylic enzymes if, for instance, in cold-adapted microorganisms, isoforms of proteins such as α-amylase or lipase exhibit high catalytic efficiency in the range of 0–30°C.

Both heat and cold exposure causes the tissue-specific expression of heat shock proteins (HSP70, HSP110, and HSP25). They function as molecular chaperones mediating the folding, assembly, or translocation across intracellular membranes of several polypeptides. Heat shock proteins also play a role in protein degradation and are required for repairing the damage resulting from stress. A subzero-tolerance mechanism, which has evolved in organisms adapted to low temperatures, is the expression of extracellular antifreeze proteins (AFPs). At least five structurally distinct AF (glyco) proteins have been identified; they all inhibit ice crystallization efficiently.

Thermosensitivity

A breakthrough in the field of peripheral thermosensation/thermonociception has been the cloning and characterization of a vanilloid receptor (VR1) present in somatic and visceral sensory neurons, trigeminal nucleus caudalis, nucleus of the solitary tract (NTS), area postrema, and caudal medulla. The cloning of another more broadly expressed receptor form (VR-like) was reported. They both probably belong to a large family of neuronal mechanosensors (responsive to stretch). VR-1 and VR-like (functional nonselective cation channels) can be activated in a narrow range of temperatures (>42°C for VR1 and >54°C for VR-L). VR1 also responds to protons and vanilloids, even when expressed in oocytes. These studies have rekindled interest in a very large number of experimental observations obtained with vanilloids (capsaicin or resinipheratoxin) that were left without molecular explanation. Capsaicin injection in the periphery does in fact stimulate sensory neurons (nociception and thermonociception) and causes the degeneration of sensory neurons after chronic treatment.

There are, however, several reports describing central effects of capsaicin on Ts neurons. Locally applied capsaicin is, for instance, able to increase the firing rate of warm-sensitive neurons and inhibit the discharge of cold-sensitive neurons in the anterior hypothalamus, causing hypothermia. Furthermore, chronic treatments can cause the degeneration of neurons in the anterior hypothalamus. Nothing is yet known about the receptor(s) that could mediate these effects centrally.

Moreover, it is not even clear whether cellular thermosensitivity is caused by the presence of a thermosensor molecule, responding to temperature changes in the same manner as acquaporins respond to osmolarity changes and opsins to light. Studies on the intrinsic thermosensitivity of different ion channels have shown that they are affected differentially by temperature changes. Among the most studied temperature-sensitive ion channels are the *Shaker*-type, voltage-gated potassium channels (Table 2).

The neurochemical features of warm sensory fibers are partially characterized. One interesting aspect, reported by several groups, is the presence of a calcitonin gene-related peptide (CGRP) in thermosensory fibers innervating, for instance, BAT. It should also be noted that cold acclimation causes a major increase in noradrenergic-, neuropeptide Y-, substance P-, and calcitonin gene-related peptide content in nerves in

Table 2 Effect of temperature on the activity (Q_{10}) of various ion channels[a]

Name	*Parameters*	*Q_{10} temperature range (°C)*	*System*	*Source*[b]
K^+channels				
Shaker H4	Activation time constant	+3.14 (20 to 4)	*Xenopus laevis* oocytes	1
	Decay time constant	+7.2 (20 to 4)	*X. laevis* occytes	1
MinK	Activation time constant	+4–7	*X. laevis* oocytes	
Na^+channels				
TTX-sensitive noninactivating Na current	Current amplitude	+4.3–7	Temperature-sensitive rat POA neurons	
Ca^{2+}channels				
HVA calcium channels	Activation time constant	+10	Dorsal raphe neurons	4
	Current amplitude	+1.7	Dorsal raphe neurons	
L-type cardiac Ca channels	Inward currents	+5.8 (15 to 25)	*X. laevis* oocytes	5
HAC		+3–4		
Cl^-channels				
From the *Torpedo* electric organ (slow gate)	Deactivation potential	+40	*X. laevis* oocytes	3
CFTR	Opening rate	+9.6		2

[a]CFTR, cystic fibrosis transmembrane conductance regulator; HAC, human articular chondrocytes; HVA, high-voltage activated; MinK, minimal potassium ion channel; POA, preoptic area; TTX, tetrodotoxin.
[b]Source:
[1]Nobile, M., Olcese, R., Toro, L. and Stefani, E. (1997). Fast inactivation of Shaker K^+ channels is highly temperature dependent. *Experimental Brain Research* **114**, 138–142.
[2]Pusch, M., Ludewig, U. and Jentsch, T. J. (1997). Temperature dependence of fast and slow gating relaxations of CIC-0 chloride channels. *Journal of General Physiology* **109(1),** 105–116, January 1997.
[3]Mathews, C. J., Tabcharani, J. A. and Hanrahan, J. W. (1998). The CFTR chloride channel: nucleotide interactions and temperature-dependent gating. *Journal of Membrane Biology* **163,** 55–66.
[4]McAllister-Williams, R. H. and Kelly, J. S. (1995). The temperature dependence of high-threshold calcium channel currents recorded from adult rat dorsal raphe neurones. *Neuropharmacology* **34,** 1479–1479.
[5]Allen, T. J. A. and Mikala, G. (1998). Effects of temperature on human L-type cardiac Ca^{2+} channels expressed in Xenopus oocytes. *Pflügers Archiv European Journal of Physiology* **436,** 238–247.

this tissue. Vanilloid receptor 1 expression and CRGP production, as well as peripheral thermosensitivity, are under the control of nerve growth factor.

Concluding Remarks

Outstanding athletic performances, obtained in the presence of prohibitive ambient conditions, are reported almost daily by mass media, and every year, as usual, animals deeply modify their metabolism to survive the coming winter. The biochemistry of the cellular/neuronal control of body temperature, however, still escapes our knowledge.

We know that autonomic and hypothalamic centers play a key role in thermoregulation, and several metabolic studies have provided a large body of evidence on the extent of the control of thyroid hormones on heat production in different organs. Further important information may come from ongoing research on the metabolic effects of leptin, from the understanding of the physiological role of uncoupling proteins, and perhaps from studies on receptors such as VR-1. The perspective looks promising, and emphasis should be placed on the multidisciplinary approach required to solve the biological problem of thermal control.

See Also the Following Articles

Heat Shock Genes, Human; Heat Shock Proteins: HSP60 Family Genes; Hyperthermia; Hypothermia; Temperature Effects; Thermal Stress; Thyroid Hormones.

Further Reading

Allman, J. (1999). *Evolving brains*. New York: Scientific American Library.

Cesare, P., Moriondo, A., Vellani, V., et al. (1999). Ion channels gated by heat. *Proceedings of the National Academy of Sciences USA* **96**, 7658–7663.

Dirig, D., Hua, X. and Yaksh, T. (1997). Temperature dependency of basal and evoked release of aminoacids and calcitonin gene-related peptide from dorsal spinal cord. *Journal of Neuroscience* **17**, 4406–4414.

Three Mile Island, Stress Effects of

A L Dougall and A Baum
University of Pittsburgh, Pittsburgh, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by A L Dougall and A Baum, volume 3, pp 595–597, © 2000, Elsevier Inc.

Stressors at Three Mile Island
Short-term Consequences
The Effects of Venting Unit 2
Long-term Consequences and Issues
Summary

Glossary

Anticipatory stress	Stress experienced as one waits and/or prepares for a stressful event. In some studies, this stress is similar to that experienced once a stressor is experienced.
Disaster	An unusual event of large scope that overwhelms the social system. Disasters can be of natural origin (e.g., earthquake, tornado) or of technological origin (e.g., Three Mile Island, war).
Stress	A psychobiological process initiated by events that threaten, harm, or challenge an organism or that exceed available coping resources. Stress is characterized by psychological, behavioral, and physiological responses that are directed toward adaptation, but can become pathological if they persist or are inappropriate.
Technological disaster	A powerful stressor that disrupts and overwhelms the social system and has a human origin, by either direct action (e.g., sabotage) or failure of technology (e.g., air disasters, industrial accidents). The human causes of these events introduce issues related to blame and control and may cause longer-lasting distress than natural disasters.
Uncertainty	A state characterized by ambiguity or lack of information pertinent to a particular situation or outcome. It is hypothesized that uncertainty is often experienced as stressful and unpleasant. Uncertainty about the extent of exposure to radiation at Three Mile Island contributed to area residents' distress.

One of the most prominent industrial disasters in United States history occurred on March 28, 1979. Failure in the core of Three Mile Island (TMI) Unit 2 resulted in the release of radioactivity into the surrounding community. For days, residents of the surrounding area, as well as people all over the nation, watched as this unprecedented nuclear accident and emergency unfolded. The accident itself was relatively brief, but uncertainty and fear about explosions and radiation release persisted for several days afterward. Authorities were unable to determine whether and to what extent residents were exposed to radiation. Poor management of the situation and conflicting reports made by authorities led to loss of trust and credibility, contributing further to uncertainty and fear among area residents. An evacuation advisory was issued for pregnant women and young children residing within 5 miles of TMI, resulting in a substantial evacuation of nearly 150 000 residents. After the initial accident and 2-week acute emergency period, the residents of the TMI area continued to deal with potentially stressful issues, such as disposal of the radioactive gas, water, and solid waste inside the reactor complex and the possible restart of the undamaged reactor at TMI (TMI Unit 1 had been shut down for routine maintenance when the accident occurred in Unit 2).

Stressors at Three Mile Island

The circumstances surrounding the disaster at TMI set it apart from other disasters such as tornados, earthquakes, and floods. Natural disasters typically cause widespread destruction of property and often involve injury or fatalities. Following the accident at TMI, however, there were no tangible negative effects, only the invisible threat of radiation and worries about its possible long-term consequences. While exposure to low-level radiation has a range of suspected effects, scientific knowledge of its effects is not well characterized. Nonetheless, many people believe that exposure to radiation causes cancers or genetic abnormalities. Area residents were uncertain about how much radiation was released during and after the accident, whether they were personally exposed to radiation, whether their exposure was sufficient to cause negative health outcomes, and whether they would develop cancer or pass on genetic defects. This occurred in the context of the loss of control suggested by analyses of human-caused disasters. The resultant stress experienced by many area residents appeared to be unusually persistent because of the continued presence of sources of new exposure and the intractability of the residents' worries about harm already done by previous exposures. The

primary stressors in the TMI environment were related to uncertainty and the life threat associated with possible consequences of radiation exposure.

Short-term Consequences

State and federal government initiatives after the accident at TMI resulted in several studies of the accident's impact on surrounding communities. In the first month after the accident, most residents reported stress and distrust of authorities, especially those residents who lived closest to the reactors, who were younger, or who were women, particularly women with preschool-aged children. Although levels of stress decreased as time passed, significant amounts of stress were still reported 9 and 12 months later, especially among residents who lived close to TMI, mothers of young children, and people who had evacuated. Residents living close to the TMI plant also reported more stress than residents living near an undamaged nuclear power plant, suggesting that it was not just the fear or distrust of nuclear power that resulted in high stress levels, but also the possible exposure to radiation following the damage at TMI.

The Effects of Venting Unit 2

While a year is usually sufficient for most people to recover from natural disasters, the events at TMI made adaptation difficult during the ensuing 12 months. Radioactive krypton gas had been trapped in the TMI Unit 2 containment building during the accident, and, beginning 16 months after the accident, officials started venting the radioactive gas directly into the environment. While the residents were assured that the amounts of radiation vented were low and safe, the venting occurred nearly every day for weeks and carried with it the threat of new contamination. Before, during, and after the venting, TMI area residents exhibited more symptoms of stress than residents living in a comparison community 80 miles away. Although there was no change in stress levels in the control sample over time, stress levels in the TMI residents decreased during and after the venting. It is possible that TMI residents were exhibiting anticipatory stress before the venting began and, once it occurred, returned to lower levels of responding. Whether this was due to an anticipatory stress response or a possible remedial effect of venting could not be addressed by available data. Regardless, TMI residents continued to experience significant stress following the venting, and by 2 years after the accident many area residents exhibited chronically elevated symptoms of stress.

Long-term Consequences and Issues

Researchers continued to follow TMI area residents during the 6-year period between the accident and restart of the undamaged TMI Unit 1 reactor. Those residents who remained in the area continued to exhibit elevated stress levels as compared to groups of people living in other communities. Negative effects were not limited to measures of self-reported distress. Heightened indicators of physiological arousal (urinary catecholamines and systolic and diastolic blood pressure), more health complaints, greater use of prescription drugs, and poorer task performance were also evident as compared to control groups. Additionally, the incident was not isolated to central Pennsylvania; people all over the country responded to the media attention given to TMI and the fears associated with use of nuclear energy and radiation exposure.

During this time, the issues surrounding the restart of the undamaged TMI Unit 1 resulted in intense legal battles that culminated in a unanimous decision from the U.S. Supreme Court to allow the TMI Unit 1 to be restarted. In October 1985, more than 6 years after the incident at Unit 2, Unit 1 was brought back online, despite protests from residents and serious concerns from community groups about the stress it would cause and the damage it would do to the healing TMI communities. Because there were no tangible environmental effects of the TMI Unit 2 accident, the possibility of psychological distress was not considered by the courts to be sufficient grounds for curtailing national efforts for nuclear energy use. However, data indicated that stress levels increased as the restart of Unit 1 approached, and, while stress levels appeared to decrease for some residents after the restart occurred, there was continued evidence for higher stress levels in area residents when compared to groups of people living in other communities. Additionally, evidence of chronic stress responding was still apparent in TMI residents up to 10 years after the Unit 2 disaster.

In addition to the factors described earlier (e.g., distance from TMI, having small children), psychosocial variables played a prominent role in the selective vulnerability of TMI residents to stress. People who reported higher perceived threat and worry over possible radiation exposure and its consequences, as well as loss of trust in officials and loss of perceived control, were also more likely to report higher levels of stress. In addition, those who felt they had fewer people to turn to for support and who tried to deny the situation or use active, problem-focused efforts to try to control an uncontrollable situation were more vulnerable to the effects of stress. In contrast, people

who tried to control their own negative emotions instead of the environment fared much better.

Summary

Although the residents of TMI suffered long-term stress and worry from the incident at TMI Unit 2, their experiences have taught stress researchers and government officials important lessons about the consequences of technological disasters. There is now greater awareness of the significant psychological and physical health effects that can occur following technological catastrophes, even when there is no tangible evidence of destruction. Perceived threat and other psychosocial mediators, such as loss of control and loss of trust in authorities, are sufficient by themselves to promote stress responding and its negative consequences. Additionally, the long-term follow-up of TMI residents has provided insight into the possible determinants of chronic stress responding. Factors such as the experience of unwanted, distressing thoughts about TMI and continued reminders of the possible danger are important determinants of vulnerability to long-term negative consequences. This information has prompted further research into these vulnerability factors as well as aided the development of interventions aimed at helping victims of other disasters.

See Also the Following Articles

Chernobyl, Stress Effects of; Community Studies; Disasters and Mass Violence, Public, Effects of; Nuclear Warfare, Threat of.

Further Reading

Baum, A. (1990). Stress, intrusive imagery, and chronic distress. *Health Psychology* **9**, 653–675.

Baum, A. and Fleming, I. (1993). Implications of psychological research on stress and technological accidents. *American Psychologist* **48**, 665–672.

Bromet, E. J. and Litcher-Kelly, L. (2002). Psychological response of mothers of young children to the Three Mile Island and Chernobyl nuclear plant accidents one decade later. In: Havenaar, J. M., Cwikel, J. G. & Bromet, E. J. (eds.) *Toxic turmoil: psychological and societal consequences of ecological disasters*, pp. 69–84 New York: Kluwer Academic/Plenum Publishers.

Bromet, E. J., Parkinson, D. K. and Dunn, L. O. (1990). Long-term mental health consequences of the accident at Three Mile Island. *International Journal of Mental Health* **19**, 48–60.

Davidson, L. M., Baum, A. and Collins, D. L. (1982). Stress and control-related problems at Three Mile Island. *Journal of Applied Social Psychology* **12**, 349–359.

Davidson, L. M., Weiss, L., O'Keefe, M. K. and Baum, A. (1991). Acute stressors and chronic stress at Three Mile Island. *Journal of Traumatic Stress* **4**, 481–493.

Dew, M. A. and Bromet, E. J. (1993). Predictors of temporal patterns of psychiatric distress during 10 years following the nuclear accident at Three Mile Island. *Social Psychiatry and Psychiatric Epidemiology* **28**, 49–55.

Dew, M. A., Bromet, E. J., Schulberg, H. C., Dunn, L. O. and Parkinson, D. K. (1987). Mental health effects of the Three Mile Island nuclear reactor restart. *American Journal of Psychiatry* **144**, 1074–1077.

Dohrenwend, B. P., Dohrenwend, S. N. and Warheit, G. J. et al. (1985). A psychophysiological field study of stress at Three Mile Island. *Psychophysiology* **22**, 175–181.

Goldsteen, K. and Martin, J. L. (1981). Stress in the community: A report to the President's Commission on the Accident at Three Mile Island. *Annals of the New York Academy of Sciences* **365**, 159–174.

Hartsough, D. M. and Savitsky, J. C. (1984). Three Mile Island: psychology and environmental policy at a crossroads. *American Psychologist* **39**, 1113–1122.

Houts, P. S., Cleary, P. D. and Hu, T. (1988). *The Three Mile Island crisis: psychological, social, and economic impacts on the surrounding population*. University Park, PA: The Pennsylvania State University Press.

MacGregor, D. (1991). Worry over technological activities and life concerns. *Risk Analysis* **11**, 315–324.

Prince-Embury, S. and Rooney, J. F. (1988). Psychological symptoms of residents in the aftermath of the Three Mile Island nuclear accident and restart. *Journal of Social Psychology* **128**, 779–790.

Thymus

M S Vacchio
National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 598–604, © 2000, Elsevier Inc.

Thymic Architecture and Components
Activities of the Thymus
Glucocorticoids and Thymic Function
Effects of Stress on the Thymus

Glossary

Negative selection	Selective inactivation of those thymocytes that express a T-cell receptor with a high avidity for self-peptide presented by self-histocompatability complex-encoded molecules and therefore are potentially harmful to the host.
Positive selection	Selective survival and expansion of those thymocytes that have successfully rearranged T-cell receptor genes and express a T-cell receptor with a moderate affinity for self-histocompatability complex-encoded molecules.
Thymocyte	An immature T cell undergoing the processes of differentiation and selection in the thymus prior to export to the peripheral immune system.

The thymus is a bilobed organ situated in the anterior superior mediastinum of mammals. The major function of the thymus is the differentiation and selection of T lymphocytes that will populate the peripheral immune system and function as part of acquired immunity. The thymus is critical for normal immune function and is highly susceptible to the effects of stress.

Thymic Architecture and Components

As early as the late 1800s it was thought that the function of the thymus was important to the immune system. However, it was not until 1961 when Jacques Miller demonstrated that thymectomy resulted in loss of graft rejection and compromised immunity to bacterial infection that this organ's place in the immune system began to be defined. The thymus gives rise to T lymphocytes, which are critical for normal immune function. Differentiation from immature precursors to mature T cells occurs in the thymus in a highly organized manner that will be discussed later in the section describing thymic function. Various aspects of thymus function have been highly characterized in mice and rats; however, the majority of these observations are also characteristic of that which occurs in humans.

The thymus consists of two main areas. The cortex normally comprises about 80% of the thymus and is densely packed with thymocytes that visually obscure the cortical stromal elements. In the medulla, the cellular density of thymocytes is considerably less and the medullary stroma is more easily visible.

The thymus differentiates from the third and fourth pharyngeal pouches in mammals and is formed from ectoderm, mesoderm, and endoderm, all of which are required for normal development. The cellular components generated by these tissues comprise the thymic stroma, composed of a diverse array of epithelial cell subsets. The origin and function of these various thymic epithelial subsets are still highly debated. Other stromal elements include macrophages and dendritic cells that are of hematopoetic origin. The cells that ultimately differentiate into T cells are derived from pluripotent hematopoetic stem cells that migrate into the thymus from either the fetal liver or bone marrow. Evaluation of the T-cell component in the adult thymus reveals significant heterogeneity. Differentiative stages of thymocytes can be determined via cell surface markers, the most characteristic of which are the expression of CD4 and CD8 and T-cell receptor (TCR) (Figure 1). The earliest thymocytes compose only a small fraction of the total T-cell number (1–5%) and express Thy-1, but do not express CD4, CD8, or TCR (double negative, DN). The population of highest frequency, composing 80–85% of thymocytes, expresses both CD4 and CD8, along with moderate levels of TCR (double positive, DP). The most mature thymocytes, poised to emigrate to the peripheral immune system as functional T cells, are the single positive (SP) cells that express either CD4 (~10%) or CD8 (~5%), along with high levels of TCR (comparable to that observed on peripheral T cells).

Activities of the Thymus

T-Cell Differentiation and Selection

The major activity of the thymus is to harbor T cells during their differentiative stages and to select functional T cells that can be exported to the peripheral

immune system. Stem cell entry occurs about gestational day 11 in the mouse (gestation is approximately 20 days). Initially, colonization occurs by penetration of the thymic capsule, although later, the influx of stem cells occurs through the high endothelial venules (HEV) that enter the thymus at the corticomedullary junction. Stem cells migrate out toward the subcapsular cortex where they give rise to the first thymocytes around gestational day 15. The earliest thymocytes that arise are the DN cells. Rearrangement of T-cell receptor genes to generate receptors of unique antigen specificity, similar to that which occurs in B cells for the generation of immunoglobulin, begins early, and by day 15 of gestation the first cells expressing the TCR$\gamma\delta$ chains ($\gamma\delta$T cells) appear. These cells migrate out of the thymus rapidly and, for the most part, are replaced with other thymocytes expressing the conventional TCR α and β chains. Rearrangement of the TCRβ chain genes significantly precedes rearrangement of TCRα chain genes, and generation of a functional rearranged and expressed β chain is required for progression to the next stage of T-cell development. Once the β chain is expressed on the cell surface as a component of the pre-TCR complex, signals associated with this expression turn off any further rearrangement of the β chain, allowing extensive proliferation and differentiation to DP thymocytes, while turning on rearrangement of the TCRα chain genes. Successful rearrangement of the TCRα locus results in moderate level expression of TCR on these DP thymocytes. It is at this point that thymocytes undergo rigorous selection processes (**Figure 1**).

One might ask why thymocytes are sequestered in the thymus during differentiation and selection. The answer to this question may reside in the manner by which T cells recognize antigen via their TCR. Unlike immunoglobulin on B cells, which bind a three-dimensional structure on an antigen, the TCR binds processed fragments of antigen that are located in a groove on self major histocompatability encoded molecules (MHC). Therefore, TCR recognition of antigen is constrained by the fact that the TCR must have affinity not only for the foreign peptide but the entire peptide/self-MHC complex. Because the specificity of the TCR is generated randomly during rearrangement of the TCR gene fragments, it is easily understandable that a large portion of T cells may generate receptors that have little or no affinity for self-MHC and therefore would be useless to the immune system. However, the specificity of a randomly generated TCR may have a very high affinity for self-peptide/MHC complexes and would be considered autoreactive and detrimental to the organism. It is unclear exactly what percentage of DP thymocytes express nonfunctional (as deemed by lack of affinity

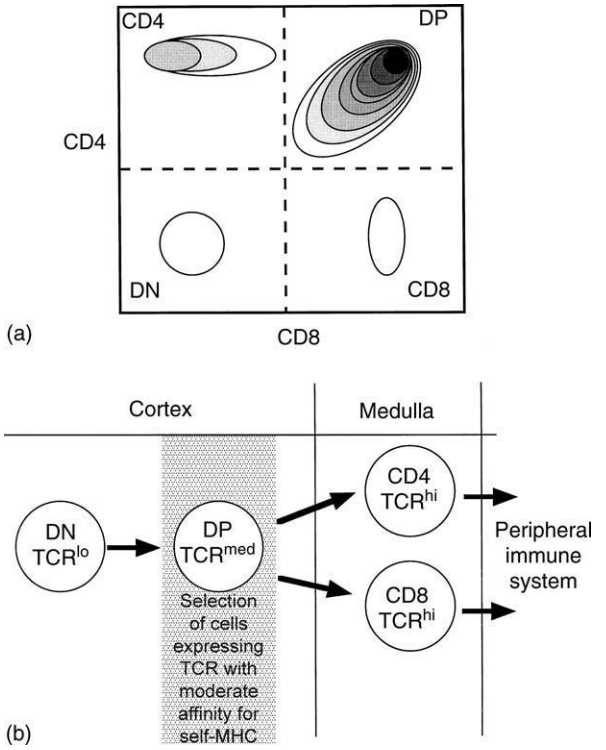


Figure 1 Thymocyte differentiation and selection. (A) Schematic diagram of thymocyte populations expressing CD4 and CD8. The earliest thymocytes (1–5%) express Thy-1 but not CD4, CD8, or TCR (double negative, DN). Double positive cells (DP) compose 80–85% of thymocytes and express both CD4 and CD8 along with moderate levels of TCR. The most mature thymocytes are the single positive (SP) cells that express either CD4 (~10%) or CD8 (~5%) along with high levels of TCR. (B) Progression and localization of thymocyte differentiation. DN cells, localized in the subcapsular cortex, differentiate to become cortical DP expressing detectable levels of TCR on their cell surface. At the DP stage, thymocytes undergo selection to eliminate cells expressing autoreactive receptors and enrich for cells expressing a TCR with moderate affinity for self-MHC. These positively selected cells differentiate further to express either CD4 or CD8 with high levels of TCR on their surface. These medullary thymocytes go on to populate the peripheral immune system.

to self-MHC) versus autoreactive TCRs, but it is clear that greater than 90% of thymocytes generated never leave the thymus. Therefore, a major function of the thymus is to cautiously screen developing thymocytes within a controlled environment to obtain those thymocytes of greatest functional potential while eliminating unnecessary clutter and the potential for damage in the peripheral immune system. It is thought that thymocytes require TCR-mediated signals to prevent programmed cell death. Thymocytes with nonfunctional TCRs do not receive this survival signal and therefore undergo apoptosis within a limited time frame. DP thymocytes in the cortex that express a TCR of high avidity for self-peptide/MHC complexes undergo a process termed negative selection, in which

elimination occurs via activation-induced cell death. These apoptotic cells are rapidly engulfed by macrophages and cleared from the thymus in a highly efficient manner so that, despite the fact that > 90% of thymocytes die in the thymus, one must look very hard to detect dying cells. Finally, those DP thymocytes that express receptors with a moderate avidity for the self-MHC undergo positive selection and differentiate into SP thymocytes with upregulated levels of TCR. Those thymocytes that recognize MHC class I differentiate to become CD8 SP cells and those that recognize MHC class II differentiate into CD4 SP cells.

Production of Thymic Factors

In addition to the role that the thymus plays in T-cell differentiation and selection, the thymus also produces a myriad of factors that can act in either a localized or a secreted manner (Table 1). Much work has focused on the thymic hormones produced by thymic epithelial cells. These were initially characterized in crude thymic extracts and have subsequently been purified and characterized. The major peptides include thymopoetin, thymic humoral factor, thymosins, and thymulin, although many less well-characterized peptides exist. The functions of these peptides are related to the maturation of thymocytes as measured *in vitro* by phenotypic changes in bone marrow after coculture and/or related to the modulation of mature T-cell function as determined postinjection of the peptide. The majority of these peptides are detectable in the circulation, and levels are influenced by a variety of factors, including thyroid hormones, pituitary hormones, prolactin, growth hormone, and stress. Thymic humoral factor has been shown to enhance lymphocyte proliferation and IL-2 production *in vitro*. Thymopoetin binds cell surface receptors on immature thymocytes and mature T cells and appears to enhance early T-cell differentiation *in vitro*. The thymosins are actually a group of peptides, the most well characterized of which are thymosin α1 and thymosin β4. Thymosin α1 induces the expression of Thy-1, CD5, and CD8 on thymocytes and enhances mitogenic, cytokine, and antibody responses of lymphocytes *in vitro*. Both thymosin α1 and β4 also increase terminal deoxynucleotidyltransferase (TdT) expression. However, of all the thymic factors, only the production of thymulin appears to be truly thymus restricted as evidenced by the observation that circulating thymulin levels decrease postthymectomy, whereas levels of the other thymic hormones are unaffected. Thymulin, which is only biologically active when complexed to zinc, is expressed by thymic epithelial cells and binds to receptors expressed on the surface of mature T cells, natural killer (NK) cells, and immature thymocytes. Functionally, thymulin can also induce expression of CD3 and CD8 on bone marrow cell and augment responses of mature T cells *in vitro*. However, while these thymic hormones can induce thymocyte changes *in vitro* or alterations in T cell function *in vivo*, a major issue that remains to be established is whether their *in vivo* function is primarily on tissues distant from the thymus or whether they also play a critical role in T-cell development *in vivo*.

Table 1 Factors produced by the thymus

Thymic hormones	*Cytokines*	*Neuroendocrine hormones*
Thymopoetin	IL-1	ACTH
Thymosin α1	IL-2	CRH
Thymosin β4	IL-4	GR
Thymic humoral factor	IL-6	PRL
Thymulin	IL-7	Somatostatin
	TNFα	Arginine vasopresin
	M-CSF	Corticosterone
	IFN-γ	

While the thymic hormones are secreted and act on other target tissues, there are numerous thymic factors that most probably act strictly in a localized manner and are potentially involved in T-cell maturation (Table 1). Cytokines are produced primarily by the lymphocytes, dendritic cells, and macrophages of the thymus. Information obtained from mice in which expression of various of these cytokines and/or their receptors was eliminated by homologous recombination has contributed significantly to mapping their relative importance to T-cell differentiation and selection. For example, while the absence of IL-2 and IL-4 appears to have relatively little effect on T-cell differentiation, IL-7 is particularly critical to progression beyond the DN stage.

Expression of neuroendocrine hormones also occurs in the thymus in cells of both epithelial and bone marrow origin. Most notably, T cells synthesize corticotropin-releasing hormone (CRH) as well as proopiomelanocortin (POMC), the products of which include adrenocorticotropic hormone (ACTH) and β-endorphin. Populations of thymic epithelial cells have been shown to express the steroidogenic enzymes, P450scc and P450cll, involved in synthesis of pregnenolone and corticosterone, respectively, and steroidogenesis occurs at levels sufficient to influence T-cell differentiation and selection. In the neuroendocrine system, CRH induces the secretion of ACTH that modulates glucocorticoid synthesis. While under normal circumstances, the production of these substances in the thymus does not occur at levels sufficient to significantly alter hypothalamic-pituitary-adrenal (HPA) axis

activity, it is interesting to speculate that the presence of these components within the thymus may lead to the paracrine regulation of localized glucocorticoid synthesis.

Glucocorticoids and Thymic Function

It has long been known that the culture of thymocytes with pharmacological levels of glucocorticoids can induce apoptosis, characterized by nuclear condensation, membrane blebbing, and DNA fragmentation. This sensitivity to glucocorticoids is the result of expression of the type II glucocorticoid receptor (GR). The GR has a 10-fold lower affinity for cortisol/corticosterone than the type I glucocorticoid receptor, also known as the mineralocorticoid receptor (MR), and is hence only activated in the presence of elevated levels of glucocorticoids. Cortical DP thymocytes are highly sensitive to exposure to glucocorticoids, whereas medullary thymocytes ($CD4^+$ and $CD8^1$ SP) and cortical DN thymocytes are relatively resistant, despite expression of the GR in all of these populations. Alterations in sensitivity can be explained by concomitant changes in Bcl-2, a mitochondrial membrane protein that can rescue cells from apoptosis. Expression of Bcl-2 is high in DN thymocytes and decreases as cells differentiate to the DP stage. However, Bcl-2 levels increase postselection and are again high in SP thymocytes, the stage at which cells become relatively resistant to corticosteroids.

While the significance of GR expression in the thymus remains unclear, recent work has shed some light on this question. Enzymes responsible for glucocorticoid synthesis have been identified in thymic epithelial cells and have been shown to be active during the neonatal period when corticosteroid synthesis by the adrenals is low. Inhibition of glucocorticoid synthesis in fetal thymic organ cultures results in abnormal T-cell differentiation and selection. Similar findings were reported in genetically engineered mice in which GR expression had been significantly decreased due to expression of the antisense GR mRNA. Analysis of the thymus in these mice revealed decreased numbers of thymocytes and decreased viability of DP thymocytes in the absence of functional GR. These data suggest that physiological levels of locally produced glucocorticoids are critical for normal T-cell differentiation and selection.

Effects of Stress on the Thymus

Involution

Exposure of mammals to changes in physical, chemical, or emotional conditions results in the generation of a "stress" response. Stress results in activation of the HPA axis, causing the release of CRH from the hypothalamus. CRH acts on the pituitary gland to facilitate synthesis and release of ACTH, which in turn induces the release of corticosterone from the adrenals. The elevation of corticosteroids in response to stress is rapid and dramatic. As discussed earlier, DP thymocytes are exquisitely sensitive to the elevation of glucocorticoids due to the expression of GR, therefore making the effects of stress on the thymus particularly evident. As early as the 1920s, Jaffe observed that the thymus involutes in response to stress and that adrenalectomy causes hypertrophy. This observation implicated adrenal products as responsible for the induction of thymic involution. However, while the correlation existed, it was not until significantly later that it was demonstrated that injection of corticosteroids mimicked stress-induced thymic involution by eliminating the majority of DP thymocytes. Futhermore, involution in stressed animals could be blocked by adrenalectomy or by treatment with the GR antagonist RU486.

Other Effects of Stress

While the most obvious effects of stress on the thymus is the dramatic decrease in size, attributable to the induction of apoptosis in DP thymocytes, glucocorticoids have a diverse range of effects other than induction of apoptosis. Elevated levels of glucocorticoids can suppress cytokines such as IL-1, IL-2, IL-3, IL-6, IL8, IL-12, G-CSF, GM-CSF, TNF-α, and IFN-γ and can induce the expression of various cytokine receptors. Some of these cytokines are expressed in the thymus and have been postulated to play roles in thymocyte differentiation. Glucocorticoid-induced downregulation of class II molecules on thymic antigen-presenting cells could prevent or hinder selection of $CD4^1$ class II-restricted T cells. Glucocorticoids released during stress are also well known to alter lymphocyte distribution, potentially due to alterations in the expression and/or affinity of adhesion molecules. Furthermore, exposure to glucocorticoids *in vitro* can induce the proliferation of thymic epithelial cells. All of these changes could significantly after thymocyte function, namely the selection of the T-cell repertoire, during stress (**Figure 2**).

In addition to increased levels of glucocorticoids, stress-induced elevation of other hormones could also potentially affect thymic function. Stress results in increased levels of ACTH, growth hormone (GH), and prolactin (PRL), receptors for which are expressed in the thymus by either thymocytes or stromal elements. The importance of GH and PRL to normal thymic function is illustrated in two strains of hypopituitary

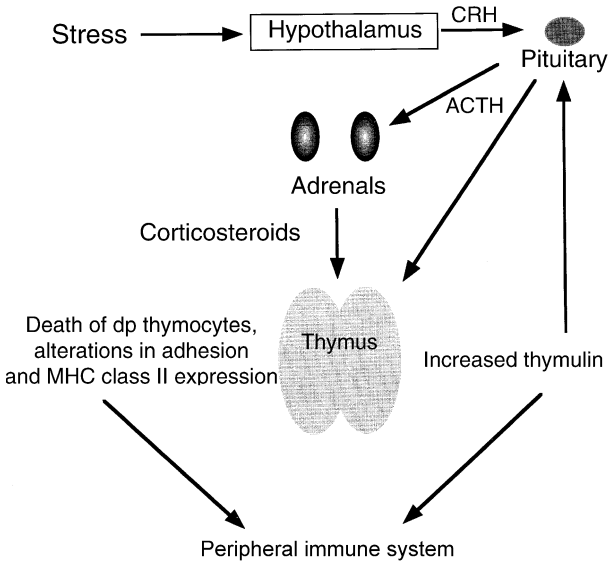


Figure 2 Schematic diagram of the effects of stress on the thymus. Stress results in activation of the HPA axis with release of CRH from the hypothalamus, thereby inducing the release of ACTH, GH, and PRL from the pituitary. The action of ACTH on the adrenals results in elevated levels of glucocorticoids that induce the death of DP thymocytes along with alterations in cell adhesion and MHC expression. Elevated ACTH, GH, and PRL can, among other potential effects, increase the production of thymulin, which can feedback on the pituitary and augment immune responses in the periphery.

mice (Snell-Bagg and Ames dwarf) that lack the acidophilic pituitary cells that secrete these hormones, resulting in small thymuses and few T cells in the periphery. These symptoms can be reversed to some degree by the transplantation of ectopic pituitary tissue. The impact of elevation of these hormones during stress has not yet been fully clarified. Both GH and PRL have been implicated independently in the promotion of thymocyte development and the augmented release of thymulin. The secretion of thymulin is also enhanced by ACTH. *In vivo* and *in vitro* studies of thymulin suggest that this peptide acts to enhance proliferation and cytokine secretion of T cells to antigen in the peripheral immune system. In addition, the secretion of thymulin may provide feedback to the HPA axis, acting on pituitary function (Figure 2). It has been shown that thymectomy results in degranulation of pituitary acidophilic cells, and implantation of neonatal thymic tissue can induce pituitary levels of lutenizing hormone and follicle-stimulating hormone.

Collectively, stress can induce multiple effects other than induction of thymocyte apoptosis. However, the contributions of many of these factors to stress-induced alterations in thymic function have been obscured by the overwhelming amount of thymocyte death during stress and have not yet been fully addressed. Stress-induced involution has been shown to be reversible in that relief of the stress results in repopulation of the thymus in mice and rats. However, because of the complexities of the system, it is not understood if there are long-term effects of stress on the thymus. For instance, despite the fact that MHC class II levels recover to normal levels post-stress, what effect does glucocorticoid-induced down-regulation of MHC class II or alteration in cytokine production and/or responsiveness during stress have on the selection of surviving thymocytes prior to recovery? Skewed selection of the repertoire would have a lasting impact on antigen responsiveness in the peripheral immune system. Nor is the functional significance of stress-induced thymic involution understood. One might speculate that in times of stress, the protection of immune status is critical and elimination of huge numbers of T-cell precursors seems, at best, counterintuitive. However, stress comes under many guises, one of which is infection by pathogens. At times in which foreign antigen is present at sufficiently high concentrations that may reach the thymus, it would be prudent to restrict thymocyte selection to avoid inadvertent tolerance induction to pathogens. One might argue further that since high levels of glucocorticoids can alter MHC class II levels and cell adhesion, selection in the thymus would be skewed compared to that which would occur during an unstressed state. Perhaps the induction of cell death in those cells actually undergoing selection, the DP cells, is the best way to prevent this potential skewing. However, this is speculation at best, and a great deal of work in this field remains before the impact of stress on the thymus and its products is fully understood.

See Also the Following Articles

Adrenocorticotropic Hormone (ACTH); Apoptosis; Beta-Endorphin; Corticotropin Releasing Factor (CRF); Immune Response; Immunity; Lymphocytes.

Further Reading

Goulding, N. J. and Flower, R. J. (1997). Glucocorticoid and the immune system. In: Buckingham, J. C., Gillies, G. E. & Cowell, A.-M. (eds.) *Stress, Stress Hormones and the Immune System*, pp. 199–224.

Hadden, J. W. (1998). Thymic endocrinology. In: McCann, S. M., Lipton, J. M., Sternberg, E. M., Chrousos, G. P., Gold, P. W. & Smith, C. G. (eds.) *Neuroimmunomodulation: Molecular Aspects, Integrative Systems, and Clinical Advances*, pp. 352–358.

Robey, E. and Fowlkes, B. J. (1994). Selective events in T cell development. *Annual Review of Immunology* **12**, 675–705.

Sprent, J. (1993). T lymphocytes and the thymus. In: Paul, W. E. (ed.) *Fundamental Immunology*, pp. 75–110.
Spangelo, B. L. and Gorospe, W. C. (1997). Thymic polypeptides and their role as mediators in neuroendocrine-immune communication. In: Buckingham, J. C., Gillies, G. E. & Cowell, A.-M. (eds.) *Stress, Stress Hormones and the Immune System*, pp. 357–372.
Vacchio, M. S., Ashwell, J. D. and King, L. B. (1998). A positive role for thymus-derived steroids in formation of the T-cell repertoire. In: McCann, S. M., Lipton, J. M., Sternberg, E. M., Chrousos, G. P., Gold, P. W. & Smith, C. G. (eds.) *Neuroimmunomodulation: Molecular Aspects, Integrative Systems, and Clinical Advances*, pp. 317–327.

Thyroid Hormones

T J Visser
Erasmus University Medical School, Rotterdam, Netherlands
E Fliers
University of Amsterdam, Netherlands

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 605–612, © 2000, Elsevier Inc.

Regulation of Thyroid Hormone Bioactivity
Effects of Stress/Glucocorticoids on Central Regulation of Thyroid Function
Effects of Stress/Glucocorticoids on Thyroid Hormone Metabolism
Conclusion

Glossary

Arcuate nucleus	Hypothalamic nucleus involved in the regulation of appetite and food intake, with important projections to the paraventricular nucleus.
Cytokines	Mediators in the response of the immune and endocrine systems to inflammation.
Deiodination	An enzymatic reaction by which iodine is removed from the thyroid hormone, resulting in the activation or inactivation of the hormone.
Glucocorticoids	Hormones from the adrenal cortex (cortisol in humans) involved in the regulation of glucose metabolism; their secretion induced by stress.
Low T_3 syndrome	The decrease in plasma bioactive thyroid hormone during nonthyroidal illness and other stress conditions.
Nonthyroidal illness (NTI)	An illness not primarily affecting the thyroid gland that is associated with a decrease in plasma bioactive thyroid hormone.
Paraventricular nucleus (PVN)	An important hypothalamic nucleus for the production of hypophysiotrophic factors.

In this article, stress is used simplistically as a general term for pathophysiological conditions associated with a stimulation of the hypothalamus-pituitary-adrenal (HPA) axis that are not caused by pathological abnormalities in the HPA axis itself. These stress conditions include those inflicted by disease, inflammation, injury, surgical trauma, fasting, heat and cold exposure, and emotional stress. In addition to stimulating of the HPA axis, these conditions are often associated with profound alterations in several thyroid-related parameters, which are commonly referred to as the low T_3 syndrome. Stress-induced alterations in thyroid hormone bioactivity result to an important extent from the centrally mediated suppression of thyroid function and the inhibition of the peripheral conversion of the prohormone thyroxine (T_4) to the active hormone triiodothyronine (T_3). Both central and peripheral effects may be mediated in part by the high cortisol levels. This article first gives a brief introduction to the regulation of thyroid hormone levels and bioactivity. Subsequently, itdescribes in some detail the stress-induced changes in thyroid hormone regulation at the central and peripheral levels and speculates about the role of increased cortisol levels in these changes.

Regulation of Thyroid Hormone Bioactivity

Normally, thyroid function is predominantly controlled by thyroid stimulating hormone (TSH), secreted by the pituitary. In turn, the TSH-producing cells of the pituitary are regulated by various hypothalamic factors, most significantly, in a positive manner, by thyrotropin releasing hormone (TRH) and, in a negative way, by somatostatin and dopamine. TSH

synthesis and release are also controlled by peripheral factors, notably by negative feedback regulation by thyroid hormone and by the inhibitory effect exerted by cortisol (Figure 1).

TRH is a tripeptide with the structure pGlu-His-Pro-NH_2 (where pGlu is pyroglutamic acid), which is produced not only in the hypothalamus but also throughout and even outside the central nervous system (e.g., in the pancreas). It is generated from a precursor protein, which in humans consists of 242 amino acids, including six TRH progenitor sequences Baa-Baa-Gln-His-Pro-Gly-Baa-Baa (where Baa is basic amino acid Arg or Lys). Enzymatic cleavage at Baa residues and further processing of the Gln-His-Pro-Gly intermediate by N-terminal cyclization and C-terminal amidation result in the generation of mature TRH.

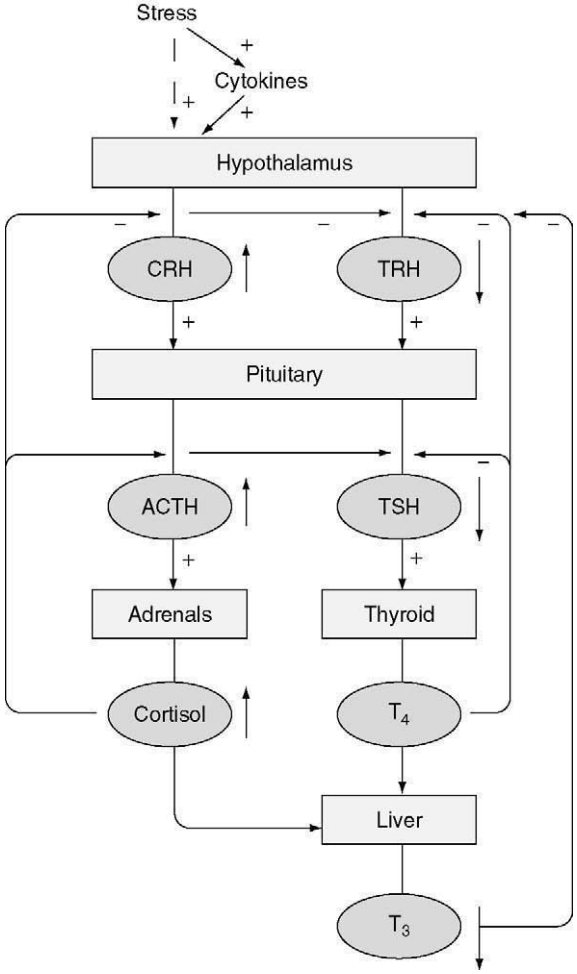


Figure 1 Stimulation of the hypothalamus-pituitary-adrenal axis by stress, the importance of cortisol in mediating the suppression of the hypothalamus-pituitary-thyroid axis, and the inhibition of peripheral T_4 to T_3 conversion. ACTH, adrenocorticotrophic hormone; CRH, corticotropin releasing hormone; T_3, triiodothyronine; T_4, thyroxine; TRH, thyrotropin releasing hormone; TSH, thyroid stimulating hormone.

Hypothalamic TRH stimulates not only TSH but also prolactin and, in some conditions, growth hormone secretion from the anterior pituitary by acting on a common (type I) TRH receptor expressed by thyrotrophs, lactotrophs, and somatotrophs. Another (type II) TRH receptor has been characterized in the brain. TSH is an ~30-kDa glycoprotein consisting of an α subunit of 92 and a β subunit of 118 amino acids. The α subunit is identical to that of luteinizing hormone and follicle-stimulating hormone, whereas the β subunit conveys hormone specificity.

TSH stimulates the production and secretion of thyroid hormone from the thyroid gland. Under normal conditions, including an adequate iodine intake, the adult human thyroid gland largely produces T_4 (3,3′,5,5′-tetraiodothyronine), which has little intrinsic bioactivity and is generally regarded as a precursor of the bioactive form of T_3 (3,3′,5-triiodothyronine). Most T_3 is produced by the outer-ring deiodination (ORD) of T_4 in peripheral tissues. Both T_4 and T_3 are inactivated by inner-ring deiodination (IRD) to 3,3′,5′-triiodothyronine (rT_3) and 3,3′-diiodothyronine, respectively. Three iodothyronine deiodinases are involved in these processes. The type I deiodinase (D1) is located in the liver, kidney, and thyroid and, surprisingly, has both ORD and IRD activities. Nevertheless, the enzyme in liver, in particular, is thought to be the main site for the peripheral production of T_3 from T_4. The type II deiodinase is expressed in the brain, the pituitary, and, perhaps also to significant extents, the human heart and skeletal muscle. It has only ORD activity and is thought to be important for the local production of T_3 in tissues such as the brain, although the enzyme in the skeletal muscle may contribute to plasma T_3 production. Finally, type III deiodinase (D3) is found in the brain, skin, placenta, pregnant uterus, and various fetal tissues. It has only IRD activity and thus catalyzes the production of rT_3 from T_4 as well as the degradation of T_3. Thus, two deiodinases (D1 and D2) are capable of thyroid hormone activation and two enzymes (D1 and D3) are capable of thyroid hormone inactivation. The relative activities of these deiodinases thus have a major impact on thyroid hormone bioavailability. The deiodinases are homologous selenoproteins, featuring essential selenocysteine residues in their catalytic centers. It is remarkable therefore that both the synthesis and the metabolism of thyroid hormone are dependent on the trace elements iodine and selenium.

The regulation of the bioactivity of thyroid hormone does not depend only on the activities of the different deiodinases; iodothyronines are also inactivated by the conjugation of the phenolic hydroxyl group with glucuronic acid or sulfate. All the enzymes

involved are located intracellularly, as are the receptors mediating the biological effects of T_3, which are located in the nucleus. Therefore, the metabolism and action of thyroid hormone in the different tissues require the uptake of T_4 and T_3 across the cell membrane. There is much evidence suggesting that the downregulation of the conversion of peripheral T_4 to T_3 during stress may be brought about by the decreased uptake and/or conversion of T_4 in T_3-producing tissues such as the liver. Because this results in decreased T_3 action and, thus, lower energy expenditure and protein catabolism, this is supposed to be a beneficial adaptation mechanism to sustain the stressful conditions.

Effects of Stress/Glucocorticoids on Central Regulation of Thyroid Function

Hypothalamic Thyrotropin Releasing Hormone

The important physiological role of hypothalamic TRH in the regulation of the pituitary-thyroid axis is clearly established. Immunocytochemical studies in the 1970s and 1980s demonstrated the presence of many (pro-)TRH-containing neurons in the PVN of the rat hypothalamus; in the 1990s the distribution of TRH cells in the human hypothalamus was reported. Although TRH-containing neurons occur in many hypothalamic and extrahypothalamic areas, TRH cells in the PVN appear to be selectively involved in the regulation of the pituitary-thyroid axis. This was first suggested by immunocytochemical studies in hypothyroid rats showing changes in TRH cells mainly in the parvocellular portion of the PVN. The assumption was confirmed by subsequent studies on TRH mRNA in the PVN of hypothyroid rats and by studies involving the endocrine effects of PVN lesions. Thus, TRH cells in the PVN are important for the set-point regulation of serum thyroid hormones, and circadian and circannual variations in TRH content, as observed in the rat PVN, may explain circadian and seasonal variations in serum concentrations of TSH.

Thyroid hormone feedback regulation is probably mediated by the thyroid hormone receptor (TR), which is expressed by TRH neurons in the PVN and by cells in the arcuate nucleus (ARC). The ARC is also one of the major hypothalamic sites expressing D2, which converts the prohormone T_4 into bioactive T_3. Studies suggest that neuropeptide Y (NPY) cells from the ARC innervate TRH cells in the PVN that project to the median eminence. Therefore, the ARC appears to be one of the pivotal loci for thyroid hormone feedback regulation in the central nervous system.

Stress, Glucocorticoids, and the Central Component of the Hypothalamic-Pituitary-Thyroid Axis

Different types of stress may induce differential changes in the central component of the hypothalamic-pituitary-thyroid (HPT) axis. For example, cold exposure in rats increases TRH mRNA in the PVN, as shown by *in situ* hybridization, and leads to increased serum concentrations of TSH. In contrast, food deprivation in rats decreases TRH mRNA in the PVN and *in vitro* TRH release and is associated with lower serum concentrations of thyroid hormones. In humans, severe illness in the absence of overt thyroid disease, NTI, is associated with a decrease in serum T_3. Despite low serum concentrations of thyroid hormones, serum TSH in NTI is typically normal or reduced. These observations point to a change in hypothalamic thyroid hormone feedback regulation. In line with this notion are profound alterations in the circadian TSH rhythm and decreased TRH mRNA expression in the PVN of patients with NTI.

Various studies have suggested a role for glucocorticoids in the alterations of the central component of the HPT axis during stress. For example, glucocorticoid excess suppresses TSH in patients and also in animal experiments. Conversely, glucocorticoid deficiency is associated with mild elevations in serum TSH, pointing to a role for glucocorticoids in the regulation of TSH release. The demonstration of glucocorticoid receptor (GR) expression by TRH cells in the rat PVN and the presence of a glucocorticoid response element in the TRH gene suggest that the inhibitory effect of glucocorticoids on TSH secretion involves the inhibition of TRH synthesis in the PVN. This suggestion was confirmed by elegant studies involving the concomitant determination of TRH mRNA and corticotropin releasing hormone (CRH) mRNA in the PVN of rats after bilateral adrenalectomy with or without subsequent treatment with glucocorticoids. After adrenalectomy, there was an expected rise in CRH mRNA, but also of TRH mRNA in the PVN. After corticosterone or dexamethasone administration, both CRH and TRH mRNA showed a major reduction. Therefore, glucocorticoids appear to inhibit TRH cells in the rat PVN, possibly directly via GR expressed by these cells. This may explain why acute bolus injections of dexamethasone suppress basal TSH release, whereas the TSH response to TRH remains intact. After long-term pharmacological glucocorticoid excess, the TSH response to TRH is also blunted. Some, but not all, studies in patients with endogenous hypercortisolism have reported an inverse correlation between parameters of cortisol excess and the TSH response to TRH, suggesting an inhibitory effect of glucocorticoids at the pituitary level (Figure 1).

Complex interactions of glucocorticoids and TRH cells have been reported in *in vitro* systems. The short-term administration of dexamethasone to dispersed hypothalamic cells showed concentration-dependent biphasic effects of dexamethasone on TRH mRNA and TRH cell content. In addition, dexamethasone stimulates TRH and TRH mRNA expression in fetal rat diencephalic neuronal cultures. In contrast to the *in vivo* situation, TRH is expressed by anterior pituitary cells in culture, which is increased further by dexamethasone. The discrepancy between the negative effects of dexamethasone *in vivo* and its positive effects *in vitro* on TRH expression in hypothalamic cells may be explained by the nonphysiological culture conditions and/or the disruption of inhibitory innervation of TRH cells in the PVN. In cultured human placental cells, cortisol at concentrations of 10–100 nmol l^{-1} inhibits D2 activity. Severe illness is associated with increased serum concentrations of cortisol. In line with decreased D2 activity in the hypothalamus of patients with severe illness is the observation that the hypothalamic concentration of T_3 (but not of T_4) is decreased markedly in postmortem tissue compared with patients with acute traumatic death. Therefore, we may speculate that the hypercortisolism of severe disease is one of the factors leading to a decreased expression of TRH in the PVN, thereby contributing to the persistence of low serum TSH (**Figure 1**). In addition, the bioavailability of T_3 may be reduced by the decreased activity of D2. In contrast, hypothalamic expression of D2 is increased in rats during starvation, again pointing to the differential effects of stress on the central component of the HPT axis according to the type of stress.

Effects of Stress/Glucocorticoids on Thyroid Hormone Metabolism

Pathophysiology of Peripheral Thyroid Hormone Metabolism

As mentioned earlier, stress conditions such as fasting, illness, and injury are associated with major changes in plasma thyroid hormone levels. Marked decreases in plasma T_3 and increases in plasma rT_3 are observed consistently in these conditions, and plasma T_4 and free thyroxine (FT_4) levels may also be affected in critical illness. In general, the magnitudes of the decrease in plasma T_3 and of the increase in plasma rT_3 are correlated with objective measures of the severity of the disease, such as loss of organ function and fever. It has been suggested that the plasma T_3/rT_3 ratio is a good indicator of the clinical condition of the patient. However, it should be mentioned that the increase in plasma rT_3 is not found in all diseases; notable exceptions are acquired immunodeficiency syndrome (AIDS) and renal disease. Initially, the stress-induced changes in plasma T_3 and rT_3 were interpreted as reflecting a shift in the peripheral metabolism of the prohormone T_4, that is, to less conversion by ORD to active T_3 and more conversion by IRD to inactive rT_3. Kinetic analysis of T_3 and rT_3 turnover during fasting and illness has confirmed that the decrease in plasma T_3 is due to a decrease in plasma T_3 production with little or no change in plasma T_3 clearance. However, the increase in plasma rT_3 seen in fasting or sick subjects appears to be caused by a decrease in plasma rT_3 clearance rather than an increase in plasma rT_3 production.

Because D1 in the liver is supposed to be an important site for the production of T_3 and the degradation of plasma rT_3, the decrease in plasma T_3 and increase in plasma rT_3 in the low T_3 syndrome may be caused by diminished hepatic D1 activity. Evidence suggests that a decrease in the hepatic content of reduced glutathione (GSH) may contribute to a decrease in D1 activity during fasting because GSH is the most abundant intracellular thiol compound and the activity of D1 is dependent on thiol cofactors. The stress-induced decrease in hepatic D1 activity may also be caused by a decrease in the expression of the enzyme at the transcriptional or posttranscriptional level. In this respect, studies of the effects of fasting on hepatic D1 expression in rats have produced confusing results. This is because fasting in rats is associated with a rapid and marked downregulation of TSH secretion, which results within 2 days in decreases in plasma T_4 and T_3 to hypothyroid levels. Strange enough, D1 gene expression is under the positive control of its own product T_3. Therefore, fasting in rats induces a hypothyroid state, which leads to a decrease in D1 expression. If plasma thyroid hormone levels in fasted rats are kept normal by T_4 replacement therapy, the decrease in hepatic D1 expression is largely prevented.

Even before D1 was identified as a selenoprotein, it was demonstrated that hepatic and renal D1 activities are strongly decreased in rats fed a selenium-deficient diet. It has subsequently been shown that a severely deficient selenium intake in humans may result in modest changes in plasma T_4, T_3, and rT_3 levels compatible with minor decreases in tissue D1 activities. Although selenium stores may diminish under certain clinical conditions, such as long-term parenteral nutrition, it is unlikely that this contributes significantly to the generation of the low T_3 syndrome. Thus, there is little solid evidence to support the hypothesis that the decrease in plasma T_3 and the increase in plasma rT_3 during fasting or other stress conditions are caused by a decreased hepatic D1 activity.

Kinetic modeling of the plasma disappearance of injected radioactive T_4, T_3, and rT_3 suggests that the fractional transfer rates of T_4 and rT_3, more than that of T_3, from plasma to the rapidly equilibrating tissue pools are decreased during fasting and illness. This is consistent with a decreased uptake of T_4 and rT_3 in the liver and, hence, a decreased intracellular availability of these substrates for conversion by D1. Studies of the uptake of different iodothyronines by isolated rat and human hepatocytes and in the isolated perfused rat liver have provided possible clues for the mechanism of these changes. These studies have produced strong evidence that the uptake of the different iodothyronines by liver cells does not occur by simple diffusion of these lipophilic compounds through the lipid bilayer of the cell membrane but is mediated by transporters that are at least partially ATP and Na^+ dependent. Competition experiments and ATP-depletion studies have suggested that uptake of T_4 and rT_3 and the uptake of T_3 by hepatocytes are mediated by different transporters, where the former is more sensitive to decreases in cellular ATP content than the latter.

The stress-induced decrease in plasma T_3 and increase in plasma rT_3 can thus be explained by a decreased activity of a hepatic T_4/rT_3 transporter with little change in the activity of the T_3 transporter. This could be the result of a decrease in cellular ATP content during fasting and illness, which would cause a much greater inhibition of hepatic uptake of T_4 and rT_3 than of T_3. Although hepatocyte incubation studies have indicated that relatively small decreases in cellular ATP suffice to cause significant inhibition of T_4/rT_3 uptake, it remains to be established whether similar diminutions in tissue ATP content may occur under pathophysiological conditions *in vivo*. Increased plasma concentrations of inhibitors are perhaps a more likely mechanism for the decreased hepatic uptake of T_4 and rT_3 during fasting and illness. Again, studies using isolated hepatocytes have provided support for this hypothesis, in that they have shown significant inhibition of T_4 uptake by various compounds, the plasma levels of which are elevated during fasting and illness (fatty acids) or specifically in liver disease (bilirubin) or kidney failure (3-carboxy-4-methyl-5-propyl-2-furan propanoic acid, CMPF; and indoxyl sulfate). The clinical relevance of these data is strongly supported by findings that in the presence of serum from patients with NTI, the uptake and metabolism of T_4 by hepatocytes are diminished compared with incubations containing serum from healthy subjects. Furthermore, the inhibition of T_4 uptake and metabolism by NTI serum are strongly correlated with the severity of the disease. Therefore, the decreased hepatic uptake of T_4 and rT_3 appears to be an important mechanism for the decreased T_3 production and rT_3 clearance during fasting and illness.

Possible Role of Glucocorticoids in the Generation of Low T_3 Syndrome

Most, if not all, conditions characterized by the low T_3 syndrome are also associated with the stimulation of the HPA axis (Figure 1). In inflammatory conditions, this is mediated by the induction of a cascade of cytokines, including tumor necrosis factor (TNF)-α, interleukin (IL)-1, and IL-6. The administration of TNF-α or IL-6 to human subjects results in a decrease in plasma T_3 and in an increase in plasma rT_3. Although these findings may reflect the indirect effects of these cytokines because their administration also induces NTI, evidence shows that IL-6 is an important mediator for the alterations in thyroid hormone metabolism. Plasma IL-6 levels generally correlate well with disease severity. Furthermore, a strong negative correlation has been observed between plasma IL-6 and T_3 levels in a number of studies in NTI patients. This correlation is improved by including the concentrations of soluble cytokine receptors.

The involvement of the cytokine network in the generation of low T_3 syndrome has also been addressed in multiple studies testing the effects of the administration of different cytokines on various thyroid parameters in rats or mice. The results of these studies have been difficult to interpret because rodents do not appear to be good models for investigation of the mechanisms by which the low T_3 syndrome is generated in humans. As mentioned earlier, the central suppression of thyroid function induced by stress is much more pronounced in rats than in humans. Moreover, transthyretin (TTR) is the main plasma thyroid hormone-binding protein in rodents, in contrast to humans in which thyroxine-binding globulin (TBG) is the major carrier. TTR is a negative acute-phase reactant; that is, its plasma concentrations fall markedly during illness, resulting in much stronger decreases in plasma T_4-binding capacity in rats than in humans. However, evidence for a possible role of IL-6 in mediating the inhibition of peripheral T_3 production during illness was obtained in mice. In different NTI models, the decrease in plasma T_3 was somewhat less in IL-6 knockout mice than in wild-type controls, whereas no effect of IL-6 deletion was seen on the decrease in plasma T_4.

Two lines of evidence strongly support an important role for glucocorticoids in the generation of the low T_3 syndrome. First, the administration of glucocorticoids such as dexamethasone and prednisone to human subjects not only inhibits TSH secretion (see earlier discussion) but also results in acute and marked decreases in plasma T_3 and increases in

plasma rT_3 levels, whereas plasma T_4 levels show little change. Second, all stress conditions associated with the low T_3 syndrome are also accompanied by large increases in plasma cortisol levels. In many, but not all, clinical studies of NTI, a strong negative correlation has been observed between plasma T_3 and cortisol levels. That such a correlation has not been found in all studies may be explained by temporal differences in the responses of plasma cortisol and T_3 to stress, with increases in cortisol preceding decreases in T_3 if they are causally related. However, a serious argument against such a causal relationship has been put forward by studies examining changes in plasma hormone levels in patients undergoing surgery under general anesthesia without or with additional epidural anesthesia. Although the surgery-induced increase in plasma cortisol is strongly diminished, but not prevented, by epidural anesthesia, the decrease in plasma T_3 and the increase in rT_3 are not affected. Although these data do not support the hypothesis that cortisol is an important mediator in the generation of the low T_3 syndrome by surgical stress, they do not exclude the possibility that the smaller increase in plasma cortisol under epidural anesthesia is still sufficient to interfere, perhaps in concert with other factors, with the peripheral metabolism of thyroid hormone.

A particularly interesting study in this respect has been conducted in rats, in which the effects of immobilization stress on plasma thyroid hormone levels were examined. This was associated with a 30-fold increase in plasma corticosterone, an $\sim 40\%$ decrease in plasma T_3, and an $\sim 75\%$ increase in plasma rT_3, as well as marked decreases in hepatic and renal D1 activities. These stress-induced changes were not seen in animals that had been surgically or chemically adrenalectomized and treated with replacement doses of corticosterone. The mechanism by which glucocorticoids affect peripheral thyroid hormone metabolism has not been established. Evidence suggests that both the uptake of thyroid hormone and D1 expression may be inhibited in the liver and kidney *in vivo*. However, *in vitro* dexamethasone has been shown to stimulate the expression of the D1 and D2 genes in various cell cultures (e.g., trophoblasts, neurons, glia cells, and hepatoma cells), perhaps by stimulating the differentiation of these cells. The physiological relevance of these observations, however, is uncertain.

It should be realized that the low T_3 syndrome may result not only from a decreased production of T_3 and clearance of rT_3 but also from an increased clearance of rT_3 and production of rT_3. It has been reported that the administration of dexamethasone to human subjects increases plasma rT_3 production, whereas plasma rT_3 clearance is not changed. Finally, it should be realized that the effects of glucocorticoids on thyroid hormone metabolism depend strongly on the stage of development. In the fetal circulation, T_3 levels are low and rT_3 levels are high. Although low expression of D1 in liver and kidney may contribute to low T_3 syndrome, the major cause is probably the high expression of D3 in the placenta, uterus, and various fetal tissues. Studies in chickens have indicated that the high D3 expression in the embryonic liver is acutely downregulated by glucocorticoids, resulting in a dramatic increase in plasma T_3. Thus, there are multiple, sometimes opposite, interactions between glucocorticoids and peripheral thyroid hormone metabolism.

Conclusion

We have discussed the decrease in TSH secretion and peripheral T_4 to T_3 conversion (low T_3 syndrome), which occur in various forms of stress. This is believed to be a beneficial defense mechanism, which saves energy and minimizes catabolism. All these conditions are also associated with a stimulation of the pituitary-adrenal axis. Together with findings that the effects of stress on plasma TSH and thyroid hormone levels are mimicked by the administration of glucocorticoids, this suggests that cortisol is an important mediator in the generation of the low T_3 syndrome during stress at both central and peripheral levels. However, studies of the effects of epidural anesthesia, which decreases the response of cortisol but not those of T_3 and rT_3 to surgical stress, suggest that there is not always a causal relationship between stress-induced changes in plasma cortisol and thyroid hormone levels.

See Also the Following Articles

Cytokines; Glucocorticoids, Effects of Stress on; Hypothalamic-Pituitary-Adrenal; Paraventricular Nucleus.

Further Reading

Bianco, A. C., Nunes, M. T., Hell, N. S., et al. (1987). The role of glucocorticoids in the stress-induced reduction of extrathyroidal 3,5,3′-triiodothyronine generation in rats. *Endocrinology* **120**, 1033–1038.

De Groot, L. J. (1999). Dangerous dogmas in medicine: the nonthyroidal illness syndrome. *Journal of Clinical Endocrinology and Metabolism* **84**, 151–164.

Docter, R., Krenning, E. P., Jong, M., et al. (1993). The sick euthyroid syndrome: changes in thyroid hormone serum parameters and hormone metabolism. *Clinical Endocrinology* **39**, 499–518.

Fliers, E., Wiersinga, W. M. and Swaab, D. F. (1998). Physiological and pathophysiological aspects of

thyrotropin-releasing hormone gene expression in the human hypothalamus. *Thyroid* **8**, 921–928.
Hennemann, G. and Visser, T. J. (1997). Thyroid hormone synthesis, plasma membrane transport and metabolism. In: Weetman, A. P. & Grossman, H. A. (eds.) *Handbook of experimental pharmacology* (vol. 128), pp. 75–117. Berlin: Springer.
Jackson, I. M. D. (1995). Thyrotropin-releasing hormone and corticotropin-releasing hormone: what's the message? *Endocrinology* **136**, 2793–2794.
Kühn, E. R., Geris, K. L., Van der Geyten, S., et al. (1998). Inhibition and activation of the thyroidal axis by the adrenal axis in vertebrates. *Comparative Biochemistry and Physiology A* **120**, 169–174.
Lechan, R. M. (1996). Functional microanatomy of the hypophysial-pituitary axis. In: Melmed, S. (ed.) *Frontiers of hormone research* (vol. 20), pp. 2–40. Basel: Karger.
Wiersinga, W. M. and Boelen, A. (1996). Thyroid hormone metabolism in nonthyroidal illness. *Current Opinion in Endocrinology & Diabetes* **3**, 422–427.

Thyrotoxicosis *See:* Autoimmunity; Graves' Disease (Thyrotoxicosis).

Tornadoes, Stress Effects of *See:* Community Studies.

Torture

I Genefke, H Marcussen and O V Rasmussen
International Rehabilitation Council for Torture Victims, Copenhagen, Denmark

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by I Genefke, H Marcussen, and O V Rasmussen, volume 3, pp 613–619, © 2000, Elsevier Inc.

Torture in the World Today
Types of Torture
Delimitation
Symptoms and Aftereffects
Rehabilitation Models
Prevention of Torture
Impunity
Conclusion

Glossary

Governmental sanctioned torture	Institutional and political torture, in contrast to torture by criminals, torture by youth gangs, domestic violence, and employer-committed harassment opposed by governmental laws.
International Rehabilitation Council for Torture Victims (IRCT)	A group founded in 1985 in Copenhagen to disseminate information about torture, its consequences, and the possibility of preventing torture. The IRCT also promotes education and training and raises international funds for the IRCT network, consisting of 98 accredited and globally distributed rehabilitation centers for victims of torture.
Rehabilitation and Research Centre for Torture Victims (RCT)	Established in 1982, a center in Copenhagen that examines and documents torture and develops diagnostics and treatment of torture victims through research.
UN Convention against Torture and Cruel, Inhuman or Degrading Treatment or Punishment	Convention (1984), with 33 articles, that defines governmental and political torture in article 1.
UN Vienna Declaration	Declaration (1993) that refers in chapter 5 to torture as "one of the most atrocious

	violations against human dignity" and "urges all states to put an immediate end to the practice of torture ... and abrogate legislation leading to impunity for torture."
World Medical Association's Tokyo Declaration	Declaration (1975) stating that "Torture is defined as the deliberate, systematic or wanton infliction of physical or mental suffering by one or more persons acting alone or on the orders of any authority, to force another person to yield information, to make a confession, or for any other reason."

Torture in the World Today

The definition of torture executed by states or state-connected institutions used in this article is that expressed by the United Nations (UN) Torture Convention, adopted in 1984 and entered into force in 1987. This convention states that torture is present when (1) severe pain or suffering, whether physical or mental, (2) is intentionally inflicted on a person (3) for such purposes as obtaining from him or her or a third person information or a confession ..., (4) at the instigation of or with the consent or acquiescence of a public official or other person acting in a official capacity. The World Health Organization (WHO) introduced in 1986 the concept of organized violence, including "torture, cruel, inhuman or degrading treatment or punishment" as in article 5 of the UN's Universal Declaration of Human Rights (1948).

Imprisonments without trial, mock executions, hostage taking, or any other form of violent deprivation of liberty fall under the category of organized violence.

Democracy and Torture

Article 8 of the Vienna Declaration states that "Democracy, development and respect for human rights and fundamental freedoms are interdependent and mutually reinforcing." From this statement, it could be concluded that a necessity for development is democracy based on the respect for human rights.

In other words, there exists a close correlation between democracy and torture. The more democracy in a given country, the less torture is carried out. The more torture, the less democracy, *inter alia,* as it appears from Amnesty International's regular annual reports.

Torture works against democracy. Torture can even be said to be a very strong tool against democracy, since torture goes against the concepts of democracy, freedom of speech and expression, and the right to political dissent. One purpose of torture is to silence the tortured and those who think alike. A result of this is that people living under pressure and lawless conditions, and consequently with prospects of violations and torture, more or less constantly find themselves in stressful situations.

Those tortured can be identified as ethnic groups and parts of a population, as individuals as well as groups, and as men and women who hold leadership positions and make a stand for democracy in authoritarian regimes. The latter may be politicians in opposition, student leaders, journalists, leaders of ethnic minorities, union members, or human rights defenders.

The Magnitude of the Problem

Torture ranks as one of the most profound human rights abuses, and the magnitude of this worldwide problem is immense. Torture as well as awareness of torture have increased. Government-sanctioned torture exists in more than 40 UN member states and is carried out sporadically in many more. A worldwide survey by Amnesty International (AI) showed that 150 countries, out of 195 investigated, practiced torture. It is also known from AI's yearly reports that torture may occur sporadically in prisons and by the police, even in civilized and democratic societies. Recently, solid evidence that torture exists even in democratic states has been demonstrated by U.S. military personnel torturing prisoners in the Abu Ghraib prison in Iraq in 2004 as well as by information of abuses and alleged torture by coalition forces in Iraq.

There is not a good methodology to calculate the magnitude of the problem of torture worldwide. The number of torture survivors is probably approximately several million.

Every year, torture forces many thousands to flee from their home country. Globally, there are more than 14.5 million refugees, and an additional 19 million have been internally displaced in their home country. A large number of these people have been tortured, although an exact figure does not exist. Community samples are rare. Jaranson and coworkers found in 2004 prevalence rates of torture ranging from 25 to 69% among selected East African communities in Minnesota. Victims of torture often do not want to report on or tell about their traumatic past, either because of fear of retaliation or persecution or because of the deep feelings of shame that are a result of the torture.

Types of Torture

The torture process typically starts with the arrest of the victim, usually at night, with a formidable display

of power and unnecessary use of violence. The softening phase, which often follows, usually consists of a couple of days and nights of unsystematic violence with beating, kicking, and other humiliations. After this, systematic torture starts when the torturers explore the weak spots of the victim to make him or her break down. The aim is to not break down the victim too quickly.

Systematic torture can be conducted in both physical and psychological forms, usually performed at the same time, which aim for a long-lasting destruction of the physical and psychological well-being of the victim. Sophisticated torture methods can cause destruction of the identity and self-respect of human beings while allowing the perpetrators to claim that the victims were never exposed to torture. A new science has developed, and modern day torture is practiced in many countries with the assistance of medical doctors and psychologists.

Physical Torture

The following is a summary of physical torture methods. Electric shocks are applied to the most sensitive areas of the body, or the victim is suspended for hours by his or her arms or by a leg. The head can be forced under water until the victim is about to suffocate, or the skin is burned by cigarettes or red-hot iron rods. The victims are beaten systematically, typically under the feet until the soles are badly damaged. Sexual offenses are common. Women in particular are attacked as sexual objects, and men are harmed in their ability to function as men. Trained dogs can be used for direct attacks or for rape of both men and women. Mock executions bring the individual to a loss of reality and a nightmarish state of almost suspended animation. The situation during the victim's detention is further worsened by filthy food and drinking water. Freedom of movement is limited, and prisoners are closely packed in small cells and thus forced to take turns sleeping. Sanitary conditions are extremely poor, and any request for visiting the toilet will often be turned into a pretext for torture.

Psychological Torture

Psychological torture takes place at the same time as physical torture: for instance, deprivation of sleep, blindfolding, and lack of human contact are forms of psychological torture. These methods leave the victim with a deep sense of helplessness, fear, and extreme stress and may cause hallucinations. The victim can be totally isolated for months or years, and during that time the victim does not know what is going to happen. The victim's family will often have no knowledge of his or her whereabouts. Psychological methods of great social and psychological impact are sexual torture and situations in which victims are forced to be present during torture to others, particularly relatives and children.

Many victims are threatened with having to do or say things that go against their ideology or religious convictions, with purpose of attacking fundamental parts of their identity, such as self-respect and self-esteem. The victims' political and ethical values are particularly attacked by the torturers when they, e.g., force the victims to sing songs that praise everything they fought against.

The attempt to break down the victims' personality often begins at the time of arrest, with the removal of personal belongings, including life-saving drugs and glasses, and replacing them with badly fitting uniforms. Names are replaced with numbers, and victims are instructed to address their jailers with great respect.

It is important to understand that the nature of torture influences even common life situations. During torture, the perpetrators often avail themselves of common objects used by everyone, such as cigarettes, telephone books, knives, needles, water, light, and noise (often the radio is tuned to music), and as a result, the survivors will in the future associate torture with these everyday objects. In such a manner the trauma of torture is specific. The survivor of torture will be reminded of his or her torture on the conscious as well as the unconscious level nearly constantly throughout the day, year in and year out. And this is what the torturer intends: that the recollection of the torture, the reminders should nag and plague the survivor for a long time or even the rest of his or her life. In the treatment process this is counteracted by psychotherapy, among other options.

Delimitation

History

The recognition of symptomatology related to torture was a phenomenon of the 1970s. At that time, there was no systematic medical literature on torture. Subsequently, a series of findings was encountered that were contrary to expectations.

The first detailed, systematic studies of the methods of torture and its immediate aftereffects were presented by Danish medical doctors in 1974–75. The search for forensic medical evidence that torture had occurred had begun. At an early stage, it was concluded that the worst sequelae of torture were psychological. That was the first surprise, confirmed by other international studies: not only is torture

unbearable and extremely painful, but it also stays with survivors and haunts them many years later.

Diagnosing Government-Sanctioned Torture

Physicians have learned much about the documentation of torture based on both physical and psychological sequelae. Examination methods, interview techniques, and data collection have become more sophisticated. Through medical examination, physicians can now detect indications of torture in most of the major organ systems, including the dermatological, cardiopulmonary, gastrointestinal, musculoskeletal, neurological, urological, gynecological, otorhinolaryngological, and ophthalmological systems, as well as in the teeth. Psychological symptoms now fit into the well-defined category of posttraumatic stress disorder (PTSD) as defined by the WHO and the American Psychiatric Association in the *Diagnostic and Statistical Manual of Mental Disorders,* 4th edition (DSM-IV). The medical profession's level of sophistication in assessing torture victims has evolved to the extent that the concept of a torture syndrome based on the results of both physical and mental torture has been proposed.

Symptoms and Aftereffects

The Effect of Torture

The mental health consequences of torture are usually more persistent and protracted than the physical aftereffects, although there is considerable overlap of the physical and psychiatric effects. Torturers today are capable of creating conditions that effectively break down the victim's personality and identity and his or her ability to live a full life with other human beings. Deep feelings of guilt and shame often occur after torture.

Feelings of guilt may be caused by the mere fact having survived when friends died while being tortured, or perhaps by having given information that could have harmed friends. Guilt may also be produced by the so-called impossible choice, when victims are forced to choose between, for instance, revealing the names of their friends and seeing members of their family tortured. Regardless of what the victim chooses, the end result is a disaster for which the victim feels responsible. And that is exactly the aim of torture as it is practiced in this day and age.

Psychological studies of survivors of torture at the Rehabilitation and Research Centre for Torture Victims (RCT) have revealed the following symptoms: anxiety, memory gaps, depression, changed personality, frequent nightmares about prison and torture, and difficulty in remembering and concentrating. Fatigue, headache, and sexual disturbances are also common. These reactions may in many situations be considered normal in ordinary people who have been exposed to something as perverse, cruel, and abnormal as torture. In situations with beatings, studies have suggested an association between head trauma and neuropsychiatric symptoms that could be recognized as cognitive deficits.

The psychological aftereffects of torture can be considered to be the same in all torture survivors, whereas the physical aftereffects of torture depend upon the type of physical torture methods used. Falanga (persistent beatings on the soles of the feet) results in impaired walking, suspension by the arms results in shoulder pain, and so on.

Over the years the examinations of many hundreds of survivors from approximately 100 countries have provided knowledge about the methods of torture and their aftereffects. These studies have made it possible to ascertain that the methods of torture used are the same all over the world. This is a natural consequence of the fact that the aim of torture is the same all over the world: to break down strong persons' identities. The methods are the same, and so are the aftereffects. This knowledge, which is most useful in medical work on torture, was obtained through systematic analysis.

Posttraumatic Stress Disorder in Torture Survivors

Neuropsychiatric symptoms are often difficult to correctly diagnose. The multiplicity of symptoms is great and comorbidity occurs. Psychiatric disorders such as major depression and PTSD are frequently seen in torture survivors. At the Centre for Victims of Torture in Minnesota, nearly 70% of the clients meet diagnostic criteria for PTSD, with nearly all of the clients exhibiting at least one or two symptoms of this disorder. Mollica and Caspi-Yavin then concluded that "the attempts to describe a single 'torture syndrome' ... are generally unconvincing." Finally, they, who have considerable experience working with refugee survivors from southeast Asia, concluded that "medical investigations of torture survivors not only failed to demonstrate a unique torture syndrome, but demonstrated symptoms closely associated with the DSM-III-R diagnosis of PTSD." Consequently, investigators have shifted their focus away from demonstrating the presence of a unique syndrome to establishing the prevalence of PTSD in torture survivors.

However, the association between PTSD and torture is not a simple linear one. Research has shown that although there is an association between torture

and PTSD, different forms of torture produce different PTSD symptoms. More specifically, individuals who experienced isolation or blindfolding, impact torture, and other types of physical torture had a predominance of PTSD intrusion symptoms, whereas individuals who had been sexually tortured described more avoidance phenomena. This may suggest that PTSD in the cases of tortured individuals is not a uniform syndrome.

It should be kept in mind that controversy exists about the applicability of diagnosing PTSD in torture survivors. In countries where torture is routinely practiced, PTSD often is considered as Western ethnocentric and limited diagnostic category that fails to capture the magnitude of torture as a trauma.

Vesti and Kastrup from the RCT concluded in 1995 that a substantial proportion of survivors developed symptomatology similar to that of PTSD but that others did not. This is consistent with the knowledge that there are individual differences in response to severe stress, which are in the early stages of being described. Basoglu et al., for instance, documented that perceived severity but not objective severity of torture was associated with PTSD, anxiety, and depression in torture victims. Impact of the captivity experience on the victim's family was the strongest predictor of PTSD symptoms. Age, personality, previous emotional and physical health, ideological and political commitment, and quality of the posttorture environment can affect the development of PTSD and other symptoms. The presence or absence of social support and the individual's perception of other persons' helpfulness are important variables for the reduction of the probability of full-blown PTSD.

Petersen et al., who reexamined 22 Greeks who had been tortured, found that eight of them met the criteria for the chronic organic psychosyndrome. These criteria included symptoms, experienced daily, of at least three of four types: (1) reduced memory or ability to concentrate; (2) disturbances of sleep; (3) emotional lability, anxiety, and depression; and (4) vegetative symptoms of the gastrointestinal or cardiopulmonary systems.

Some centers in developing countries placed an emphasis on the recognition of torture as a medical issue and the need for a formal account of the symptoms that follow severe violations of human rights. These centers pointed out that no compensation is currently given to victims and that there is no understanding in their populations or among their politicians that sequelae of torture exist. Therefore, the diagnosis of posttorture syndrome is important.

Regarding the existence of a torture syndrome, the International Rehabilitation Council for Torture Victims (IRCT) acknowledges that influential researchers on torture have shifted their focus away from a unique torture syndrome and includes the psychopathological syndromes following torture as a subtype of the PTSD category. However, the criteria for PTSD are not sufficient for the categorization of the entire picture after torture. The psychological and physical profiles of PTSD and the posttorture state diverge considerably.

Symptoms that are core criteria for the definition of the posttorture state are described in the DSM-IV definition of PTSD as associated symptoms that are not necessary for a diagnosis of PTSD. These core symptoms are survivor's guilt with low self-esteem, changed personality (the continuity, wholeness, and autonomy of the self is negatively affected), physical sequelae without organic substrate, and many physical complaints without corresponding medical findings. Diffuse pain is a crucial element in the physical and psychological profile of the torture survivor.

Further research in this field might suggest that when the extreme traumatic stressor is interpersonal, the posttorture psychological and physical picture cannot be adequately described by the PTSD entity as it is currently defined. Systematic torture constitutes a more fundamental assault on the individual's self and assumptive world than random violence or other forms of extreme trauma, such as natural disasters. Therefore, there is a continuing need for specialized research on the hypothesis of a torture syndrome.

Rehabilitation Models

Today we know how to diagnose and document torture. The basic principles of diagnosis and documentation are given in the Istanbul Protocol, a UN manual on the effective investigation of torture. Also, today we have the knowledge of how to rehabilitate torture victims, and several different rehabilitation models have been developed.

The rehabilitation model most often used is the holistic approach, in which the psychological, somatic, social, legal, spiritual, family, and cultural aspects are taken into consideration. A point of conceptual importance is to consider torture survivors as not having any premorbid psychopathology.

The treatment principles of this approach are (1) to treat physical and psychological symptoms at the same time, (2) to secure the patient's trust and confidence, (3) to respect the individual, (4) to avoid situations that remind the patient of torture, and (5) to inform the patient carefully about examinations. Physical treatment in relation to diagnostic examination can be practiced, and fair results are possible even many years after torture. Repair of pain

is a very important goal for treatment, and the medical profession can in many cases accomplish that goal.

Treatment rehabilitation methods can be individual, couple, family, or group therapy or counseling, which is a method that, to some degree, takes into account a deeper understanding of the individual's experiences, giving relief and understanding and helping the person to solve problems and make choices.

Other forms of therapy include the following: psychological insight therapy, used for the meeting, the initial setting, the emotional phase, reintegration, and the end of therapy; psychological supportive therapy, used to deal with practical matters such as bodily dysfunctions, social matters such as housing and language, and integration matters, for refugees to get a basic understanding of their new country, as well as to provide balance between insight and supportive psychotherapy; special programs for children and families, such as couples therapy, family therapy, individual child therapy, group therapy, and network meetings; and the outreach method for screening and rehabilitation of torture victims. This last model has four points of importance: availability, accessibility, adaptability, and appropriateness. Availability refers to the services being available to meet the victims' health needs, both general and special. Accessibility means that the services must not be too distant or culturally insensitive. Adaptability refers to the likelihood of the project being acceptable to the refugees and therapeutically effective at the same time. Appropriateness is related to the question of whether a project is appropriate to the refugee groups covered by it.

The aim of all rehabilitation models is to facilitate the social functioning of the victim by strengthening adaptive coping mechanisms, discouraging nonadaptive processes, and facilitating access to social opportunities. Furthermore, it aims to facilitate the psychological treatment of traumatic experiences, loss, and bereavement by consoling, comforting, and protecting victims, sharing their experiences and emotions, recognizing their suffering and pain, and supporting them in their expression of grief when the first survival needs are satisfied. The final step is to treat the victims who have developed psychological disturbances.

Counseling is defined as a method that, to some degree, takes into account a deeper understanding of the individual's experiences, providing relief and understanding and helping the person to solve problems and to make choices. The counselor must respect the confidentiality of information given and must be capable of handling situations that are loaded with painful feelings without getting lost, and using them for the benefit of the client. He or she must also be skilled in analyzing problems, giving advice, and knowing how to help the client arrive at his or her own decisions.

The counseling program is aimed at recruiting persons who are not health professionals such as doctors or psychologists to act as counselors. The program was created because in many countries health professionals are not available in sufficient numbers. However, other very capable professionals, such as teachers, journalists, monks, and nuns, will be able to learn the program and do a good job as counselors – not necessarily at the same level as the trained psychotherapist, but they will still be able to provide valuable help.

The aim of counseling is to increase awareness of experienced traumas and the reactions to them. Furthermore, it aims to provide relief from psychological suffering, to reduce aftereffects of traumatic experiences, and to help the victim regain control of life situations in order to become, once more, a fully integrated member of society.

The majority of torture victims cope with the help of family or community. Some of them, because of the severity of their symptoms, are not able to reintegrate into society and thus need assistance. Assuming the efficacy of treatment, an important challenge is to identify these victims and help in their communities.

Prevention of Torture

Prevention of torture can be differentiated into the steps of primary prevention, directed toward society itself, secondary prevention, directed toward groups within society, and tertiary prevention, directed toward individuals. Primary prevention of torture involves creating awareness of the problem by making torture visible. This can be done by establishing centers and programs for the rehabilitation of victims of torture. Information and education are also very important tools in the endeavor to create awareness.

Secondary prevention is directed toward teaching relevant personnel – military, police, and prison personnel – that torture always is prohibited, even when an order to torture is given.

The rehabilitation of torture victims serves as tertiary prevention against further violations involving torture, as the presence of rehabilitation services renders the problem visible. Torture may induce a belief in violence and loss of confidence in humanity, but rehabilitation can make people believe in humanity again. In the long run, rehabilitation can help the development of democracy and prevent future conflicts in favor of less stressful conditions.

The UN Convention against Torture contains all the necessary provisions to be able to punish perpetrators and to compensate, prevent, control, educate,

and inspect. This instrument was adopted in 1984 and came into force on June 26, 1987, which was proclaimed as the UN International Day in Support of Victims of Torture – a day not only for creating awareness but also for providing moral support to torture victims. It is a symbolic gesture that recognizes the victims and their families and honors those who did not survive the torture. The convention has been ratified by 141 countries, including all of those countries that had troops in Iraq in 2005 and until now.

Support from the UN, many governments, human rights organizations, nongovernmental organizations, and numerous initiatives is important in helping to break through the silence, insecurity, and indifference surrounding the issue of torture and will help make a stand for visibility, for openness, for acceptance among the boards of various foundations, and for the understanding of the necessity for moral rehabilitation of torture victims.

Impunity

The UN Convention states that torture is an offense under the criminal law of the country in which it occurs and that persons who are accused of torture are to be taken to court and, if found guilty, given penalties that take into account the grave nature of the offense. Article 60 of the Vienna Declaration points to the problem of impunity. It states that legislation leading to impunity for those responsible for torture and other grave violations of human rights should be abrogated.

Not much progress has been made in this area, although promising changes in public opinion in Chile and Argentina have led to ongoing legal proceedings against tyrant generals from the time of dictatorships. In several countries where torture takes place or has taken place, the identity of many of the torturers and those responsible for torture is well-known. Still, there are very few examples of these persons being punished, which is problematic for several reasons. First of all, this is troublesome for the moral rehabilitation of the victims of the torture. If the torturers are not punished, the victims will easily be left with the sentiment of being powerless and humiliated. Second, victims of torture go through unnecessary fear and distress when exposed to the risk of encountering their torturers in the streets. Third, impunity creates a twisted set of norms and values in society. Allowing the persons responsible for gross human rights violations to go free while at the same time hitting hard on conventional crime can lead to an increased brutalization of society. The veil of impunity is broken down slowly but surely, and medical professionals, with their ability to diagnose and document the incidence of torture, have a key role to play in ensuring that torturers face prosecution and that the victims of torture receive proper rehabilitation and compensation.

The work of the Criminal Tribunal in the Hague should be mentioned. The work of the tribunal has shown the need for the establishment of a permanent international criminal court. This initiative can be a future platform for prosecution of those responsible for torture and thereby an important instrument for the eradication of torture and the moral rehabilitation of victims of torture. It is of great importance, however, that the protection and the support of the victims will be ensured by this court.

Conclusion

More than 30 years of professional, medical, and psychological work on torture have taken place. Today we may say that we have enough knowledge about the phenomenon of torture and enough conventions and declarations. Also, the experience available makes it possible to talk about torture and its effects in a more substantial and assured way. Sufficient basic knowledge (i.e., social analysis on a medical and psychological basis) allows us to make strong statements against the practice of torture. What is needed, however, is national and international recognition of the phenomenon of torture as well as recognition of the magnitude of the problem and its transgenerational effects and the necessity of implementation of conventions and declarations.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Posttraumatic Stress Disorder – Clinical; Posttraumatic Therapy.

Further Reading

Allodi, F. (1991). Assessment and treatment of torture victims: a critical review. *Journal of Nervous and Mental Disease* **179**, 4–11.

Amnesty International. (2000). *Torture worldwide. An affront to human dignity.* New York: Amnesty International USA.

Basoglu, M. (ed.) (1993). *Torture and its consequences.* New York: Cambridge University Press.

Basoglu, M., Jaranson, J. M., et al. (2001). Torture and mental health: a research overview. In: Gerrity, E., Keane, T. M. & Tuma, F. (eds.) *The mental health. Consequences of torture*, pp. 35–62. New York: Kluwer Academic/Plenum.

Dunér, B. (ed.) (1998). *An end to torture.* London: Zed Books.

Elsass, P. (1997). *Treating victims of torture and violence. Theoretical, cross cultural, and clinical implications*. New York: New York University Press.

Hougen, H. P., Kelstrup, J., Petersson, H. D., et al. (1988). Sequelae to torture: a controlled study of torture victims living in exile. *Forensic Science International* **36**, 153–160.

Iacopino, V., Ozkalipci, O. and Schlar, C. (1999). The Istanbul Protocol: international standards for the effective investigation and documentation of torture and ill treatment. *Lancet* **354**, 1117.

Jaranson, M. and Kastrup, M. (2004). Developments and progress, 2001–2004. In: World Psychiatric Association (ed.) *Section on the psychological consequences of torture and persecution*. Lavis, Italy: Masson.

Mollica, R. F. and Caspi-Yavin, Y. (1992). Overview: the assessment and diagnosis of torture events and symptoms. In: Basoglu, M. (ed.) *Torture and its consequences: current treatment approaches*, pp. 253–274. New York: Cambridge University Press.

Peel, M. and Iacopino, V. (2002). *The medical documentation of torture*. London: Greenwich Medical Media Ltd.

Petersen, H. D., Abildgaard, U., Daugaard, G., et al. (1985). Psychological and long-term effects of torture: a follow-up examination of 22 Greek persons exposed to torture, 1967–1974. *Scandinavian Journal of Social Medicine* **13**, 89–93.

Quiroga, J. and Jaranson, J. M. (2005). Politically-motivated torture and its survivors: a desk study review of the literature. *Torture* **15**, 1–112.

Rasmussen, O. V. (1990). Medical aspects of torture: torture types and their relation to symptoms and lesions in 200 victims, followed by a description of the medical profession in relation to torture. *Danish Medical Bulletin* **37**(Supplement 1), 1–88.

Somnier, F. E. and Genefke, I. K. (1986). Psychotherapy for victims of torture. *British Journal of Psychiatry* **149**, 323–329.

Torture, quarterly journal on rehabilitation of torture victims and prevention of torture, 1991–2005.

Vesti, P. and Kastrup, M. (1995). Refugee status, torture and adjustment. In: Freedy, J. R. (ed.) *Traumatic stress: from theory to practice*, pp. 213–235. London: Plenum Press.

Transcortin *See:* Corticosteroid-Binding Globulin (Transcortin).

Transport-Related Stress

R G Smart
Centre for Addiction and Mental Health, Toronto, Canada

© 2007 Elsevier Inc. All rights reserved.

The Natural Stress of Driving
Driver Stress and Driver Behavior
Stress after Motor Vehicle Accidents
Aggressive Driving and Stress
Driving Stress and Road Rage
What Can Be Done to Reduce Stress in Driving?

Glossary

Acute stress disorder	Severe stress resulting from a dangerous situation and lasting up to 4 weeks.
Driver stress	Stress that is not trait stress but relates specifically to the problems and challenge of driving.
Posttraumatic stress	Severe stress resulting from a dangerous situation, marked by anxiety, threat, fear, and helplessness; role impairment; and a preoccupation with the situation. This stress lasts more than 4 weeks.
Road rage	Situations in which drivers or passengers attempt to kill, injure, or intimidate another driver, another passenger, or a pedestrian.
Trait stress	Stress that is not situational but is related to ongoing personality and lifestyle factors and is an enduring feature of a person's life.

The Natural Stress of Driving

Advertisements for new cars often show drivers racing along an open road, luxuriating in the comfort and phenomenal speed of the car. There are no other cars about, no backseat drivers, and no children

crying, "Are we there yet?" The reality of driving can be very different. Driving is often done in small cars with restricted space and difficult passengers. The environment is frequently noisy because of horns honking and the din of traffic. Attentional overload because of the complexity of modern cars may also create stress. There are often time restrictions: getting to work on time, picking up children, or making it to appointments. Many roads are congested, under repair, or closed altogether. There are strict performance demands – we must pay attention, avoid accidents, and cope with some other drivers' wrath. Driving in heavy rain and in the winter in northern countries creates more stress.

Drivers are usually immobile and cannot get out, especially on freeways. Driving can also be dangerous because traffic lights change quickly and other drivers tailgate, cut in and out, or express their anger in other ways.

Many roads in or around cities are congested, and the levels of congestion are increasing in most North American cities. Highway building has not kept pace with car ownership. In North America, there are, on average, two vehicles per family, whereas there was only one vehicle per family for the previous generation. People are spending longer times in their cars and making longer commutes to work or school. Hence, they expose themselves to greater stress on the road.

Driving is often seen as stressful and difficult. In a recent Canadian study, 55% of drivers said that some of their driving was stressful and 15.7% said that half or more was stressful. Approximately 75% said that half or more of their driving was on busy roads and approximately 50% said that most or all of it was on busy roads. Considerable research shows that most drivers find driving on congested roads to be more stressful.

Some drivers have physical limitations that make driving more stressful for them. These include people with arthritis, with neurological or serious psychological problems, and who are elderly or disabled. Cars are not well designed to suit the needs of such people, and driving demands are difficult for them to fulfill. Elderly people often stop driving for physical or psychological reasons. However, the loss of driving capabilities often causes more stress and can lead to depression and social isolation for the elderly.

Driver Stress and Driver Behavior

Particular driving events cause stress among drivers, but many drivers have high levels of stress before they get in the car. Trait stress, or ongoing nonsituational stress, has been much studied and several scales have been developed to measure it. Driver stress is greater in situations in which there is a lot of congestion and in which there is time urgency. Daily hassles contribute to driver stress. Drivers who recently had an illness, personal conflict, or bereavement are more likely to be in serious accidents.

Total driver stress varies because of the driving situation, personal disposition or trait stress levels, and nondriving events such as daily hassles. High levels of stress in drivers cause attention lapses, errors, and traffic violations. In several studies, driver stress has also been related to aggression and to involvement in accidents. However, most driver stress dissipates quickly once the driver gets out of the car.

Several methods of reducing driver stress have been investigated. Commuters who worked at a job with flextime had less driver stress than those who always traveled during rush hours. Drivers on flextime also felt less time urgency. Experimental studies have shown that drivers who viewed computerized highway drives that had more vegetation than human-made structures had less stress. Also, listening to music may reduce driver stress.

Stress after Motor Vehicle Accidents

Many drivers experience long-term stress after an injury in a motor vehicle accident (MVA). This stress can last many months or even years. The proportion of MVA victims having serious postaccident stress varies from study to study. Accident-induced stress is often assessed using the criteria from the *Diagnostic and Statistical Manual of Mental Disorders*. Acute stress disorder (ASD) is diagnosed within 1 month of the accident, and posttraumatic stress disorder (PTSD) is not diagnosed until after stress has lasted more than 1 month. Both involve subjective feelings of anxiety, fear, and helplessness; role impairment; and a preoccupation with the accident and its consequences.

Most studies of ASD and PTSD involve victims who were injured in accidents and assessed in emergency wards. In such samples, the rate of ASD varies from 10 to 40% and the rate of PTSD from 10 to more than 70%. The highest rates have typically been found in children injured in traffic accidents and people involved in litigation after an accident. Usually, victims with ASD are far more likely than others to develop the long-term stress syndromes seen in PTSD. However, some victims do not develop PTSD after having ASD. Most, but not all, studies found little correlation between PTSD and the severity of the injury.

In children, the likelihood of PTSD often relates to the seriousness that they attribute to the accident. Children after accidents are often afflicted with anger, alienation, thought suppression, and intensive memories of the accident. Most studies show that PTSD is much more common among female than male victims. PTSD may be also more common among those with lower socioeconomic status.

Postaccident stress is not found only among the victims of accidents. Parents also experience considerable stress if their children have PTSD. Also, PTSD, guilt and shame have been found among drivers who caused deaths in motor vehicle accidents. In that group PTSD and guilt were associated with the degree of responsibility the driver assumes for the accident and the severity of their punishment.

Post accident stress can last a long time and can also be related to other negative life events. Among PTSD victims, events such as loss of job or income, broken relationships, and serious illnesses were more common than among accident victims without PTSD. Because of legal liability and compensation issues, some postaccident stress may be faked or exaggerated. Some studies have shown that actors can fool evaluators of stress, especially if little corroborative evidence is available.

Aggressive Driving and Stress

Aggressive driving often causes stress for other drivers as well as for the aggressive driver. Aggressive driving typically involves tailgating; cutting in and out of traffic; preventing other drivers from passing; not moving forward on green lights; and shouting, gesticulating, or criticizing other drivers. If there is an attempt to threaten or intimidate other drivers or to damage their vehicle, this is referred to as road rage, but the two terms often overlap.

Angry or aggressive drivers are more likely to be male and young. Scales for measuring trait driver anger have been devised. Drivers who score high on these scales have more anger and anxiety and show less control over their aggression; they have more minor accidents and close calls but probably not more major accidents.

Driver anger is common and probably comes from feelings of stress during driving. Approximately one in four college student drivers reported being angry with another driver once or more per day. They rated stress from other drivers as equal to the stress of college examinations; traffic congestion, road construction, and finding a parking place were seen as less stressful.

Driving Stress and Road Rage

Road rage may appear to be a new cause of stress for drivers. However, it has been around for a long time. Sophocles in his play *Oedipus the King*, written in about 420 BCE, made road rage the reason that Oedipus kills his father. Also, the poet Byron was engaged in several incidents of road rage in the early 1700s. There are other historical cases as well. It is clear that mass media reports of road rage have recently increased in several countries, although one survey shows that incidents of road rage have been decreasing recently.

Road rage has no accepted definition; it has been defined as a situation in which a driver or passenger attempts to kill, injure, or intimidate another driver, another passenger or a pedestrian in a driving incident. Cases have been described in many countries in North America and Europe, and many have led to death or serious injury. A national study in the United States found that 30% of respondents had complained about other drivers; 17% had yelled at other drivers; 3% had chased other drivers; and 1–2% had gotten out of their cars to hurt other drivers, deliberately hit other drivers, or had carried a weapon. A study in Arizona found that 28% of respondents had aggressively blocked other drivers or followed them to retaliate. Approximately 11% always (4%) or sometimes (7%) carried a gun in their car, and hostile driving behavior was much more common for those with guns.

A Canadian study found that almost half of drivers each year were cursed at or had gestures made to them. Approximately 7.2% were threatened with personal injury or damage to their cars. This study and several others found that perpetrators of road rage are more likely to be younger and male. The victims and perpetrators of road rage are often the same people. However, there is a group of frequent perpetrators of road rage that accounts for almost all of the serious cases involving injury.

Road rage can be a very stressful event for its victims, especially if there is an injury. These victims may remember the event for long periods, modifying their driving behaviors to avoid later incidents. Victims of road rage often feel trapped and unable to get to safety. The immediate effects of road rage are fear, anxiety, anger, and annoyance. Victims report being shaken by it and sometimes being confused. Some victims have lingering thoughts and negative emotions after road rage, but most of this disappears in a few weeks unless there is a serious injury. Being the victim of road rage does make some drivers more cautious about driving.

Road rage is often associated with psychiatric morbidity and distress. In several studies, frequent perpetrators of road rage were found to have high rates of psychiatric distress and, more significant, stressful life events. It may be that frequent perpetrators of road rage bring their stress into the driving situation and, because of their high trait stress, are easily upset in traffic, lashing out at other drivers frequently.

Some studies have also shown that victims of road rage have high levels of psychiatric distress (i.e., anxiety, depression, and stress). Being a victim in a serious incident of road rage is likely to be traumatic for many drivers. The direct casual relationship between serious, injurious incidents of road rage and stress levels has not been examined, but such a relationship is highly plausible given what we know about stress after traffic accidents.

Road rage has been found to be more common among drivers who have more stressful driving experiences. Studies in Canada found that road rage was more frequent for drivers who often drove in congested driving conditions, whose driving was mostly done on busy roads, and who drove longer distances. Road rage was also more common for truck drivers; they drive in stressful situations with tight time schedules, and they are often tired.

Bus drivers have been found in several studies to be both the victims and perpetrators of road rage. They are often victimized by angry passengers, and such conflicts often lead to burnout and the desire to leave the job.

What Can Be Done to Reduce Stress in Driving?

Some of the stress drivers experience is trait stress that they had when they entered the car, and some comes from the experiences they have when driving. Reducing trait stress may require drivers to change relationships, work, and life styles and get better control over their lives. Reviewing these changes is beyond our scope here and we address only the stress that is directly related to driving.

Much stress in driving relates to traffic congestion, and most large cities are experiencing greater congestion. Fully automated transportation systems may eventually divert traffic from congested areas. A few cities are attempting to reduce congestion by increasing public transit, limiting rush-hour traffic, charging motorists for using certain roadways, and the like. Whether and how these measures reduce stress to reasonable levels is unclear at this time. More could be done to allow flextime and reduce rush-hour travel. Roads could be built to be less complex for drivers, perhaps through traffic calming or other engineering changes.

Much of the serious stress from driving derives from involvement in accidents, especially those with injuries. In North America, the rates of fatal and injury-causing accidents have been declining for many years. The reasons for these declines are the greater use of seat belts and air bags, reduction in alcohol-related accidents because of strict enforcement, better licensing programs for young drivers, and engineering changes to make car crashes more survivable. The extension of these policy changes to other geographic areas could further reduce serious accident levels; so could more technical developments in cars and highways to make them safer.

Some driving stress comes from road rage and aggressive driving. Methods for reducing road rage include better education programs and anger management for convicted drivers as well as changes to vehicles. Some vehicle manufacturers such as Jaguar and Mercedes make cars that have radar systems to prevent tailgating, a common cause of road rage. These systems could be made more available. Cars could be redesigned to have voice-activated warnings when drivers raise their voices. Also cars could have lights that signal when drivers are becoming victims of road rage. Further modifications could also be made to track the perpetrators of road rage with global positioning systems (GPSs) and provide warnings to them. Much driver aggression relates to traffic congestion, and anything that reduces it will lower road rage, aggressive driving, and stress levels for drivers.

Acknowledgment

This research was supported by a grant from Auto21, one of the Networks of Centres of Excellence.

Further Reading

Blanchard, E. B. and Hickling, E. J. (2004). *After the crash: psychological assessment and treatment of survivors of motor vehicle accidents*. Washington, DC: American Psychological Association.

Hancock, P. A. and Desmond, P. A. (2001). *Stress, workload and fatigue: human factors in transportation*. Mahwah, NJ: Lawrence Erlbaum.

Hennessy, D. A. and Wiesenthal, D. L. (2001). Gender, driver aggression and other driver violence: an applied evaluation. *Sex Roles* **44**(11–12), 661–676.

Hennessy, D. A., Wiesenthal, D. L. and Kohn, P. M. (2000). The influence of traffic congestion, daily hassles and trait

stress susceptibility on state driver stress: an interactive perspective. *Journal of Applied Biobehavioural Research* **5**(2), 162–179.

Lowinger, T. and Salomon, Z. (2004). PTSD, guilt and shame among reckless drivers. *Journal of Loss and Trauma* **9**(4), 327–344.

Mather, F. J., Tate, R. L. and Hannon, T. J. (2003). Post-traumatic stress disorder in children following road traffic accidents: a comparison of those with and without mild traumatic brain injury. *Brain Injury* **17**(12), 1077–1087.

Mizell, L. (1997). *Aggressive driving*. Washington, DC: Automobile Association for Traffic Safety.

Smart, R. G., Asbridge, M., Mann, R. E., et al. (2003). Psychiatric distress among road rage victims and perpetrators. *Canadian Journal of Psychiatry* **48**(10), 681–688.

Smart, R. G. and Mann, R. E. (2002). Is road rage a serious traffic problem? *Traffic Injury Research* **3**(3), 183–189.

Smart, R. G., Mann, R. E., Zhao, J., et al. (2005). Is road rage increasing: results of a repeated survey. *Journal of Safety Research* **36**(1), 195–201.

Trans-sexualism

R A Allison
American Medical Association, Phoenix, AZ, USA

© 2007 Elsevier Inc. All rights reserved.

Introduction and Definitions
Childhood Experiences of Transsexual People
Factors that Discourage Transition
The Transition Process
Hormonal and Surgical Changes
Health Effects of Stress for Transsexual People
Religion and the Transsexual Person
Violence and Fear
Intimacy and the Transsexual Person
Aging and Death and the Transsexual Person
Conclusion

Glossary

Assignment	The process society follows to designate a person as a male or a female.
Crossdresser	A person who wears the clothing of the opposite sex, but does not self-identify as a transsexual.
Female-to-male (FTM)	A person born biologically female, whose gender identity is male, who may undergo medical and surgical changes to confirm her male identity.
Gender dysphoria	A person's persistent feeling of discomfort with the gender assigned at birth.
Gender identity	A person's inner feeling of self-identification as male or female. Gender identity in some people may be intermediate, neither fully male nor fully female. Some people experience changes in their understanding and acceptance of gender identity with the passing of time.
Gender identity disorder (GID)	A medical term, used as a synonym for transsexualism. GID is listed in the *Diagnostic and Statistical Manual of Mental Disorders*.
Gender role	The category (male, female, or ambiguous) in which society places a person, based on physical characteristics and behavior.
Male-to-female (MTF)	A person born biologically male, whose gender identity is female, who may undergo medical and surgical changes to confirm her female identity.
Reassignment	Actions taken to change society's designation of a person as a male or a female.
Transgender people	Transsexual people, crossdressers, and people of less-fixed gender identity; sometimes used to describe people who live in their desired gender role but do not seek surgical reassignment.
Transition	The process a person follows to live in his or her preferred gender; encompasses the physical changes brought about by hormones and surgical procedures, as well as the social changes of experiencing life and relationships in the new gender.
Transsexual	A person who experiences persistent and severe discomfort in the gender assigned at birth and who wishes to permanently live in the opposite gender role, with all the physical and social changes that role implies.
Transvestite	A crossdresser. This is an older term, used in the mid-twentieth century, and is not currently preferred by most people because it has been associated with a strictly sexual motivation.

Introduction and Definitions

We humans choose certain characteristics by which we define ourselves. The earliest and most essential definition is whether an individual is a man or woman, boy baby or girl baby. The distinction is established at birth, and nurseries are furnished in pink or blue. The basis for the distinction is our external appearance. Our genitalia determine the life we are expected to lead. All humankind is divided distinctly into these two groups: except when it is not.

In fewer than 1/1000 children – the true incidence is still not known – a persistent discomfort exists with the individual's physical sex. This discomfort is present from earliest memory. There is no identifiable behavioral influence to produce it. The child grows and matures unremarkably, to all external appearances; but the child's thought patterns and behavioral instincts are those of the opposite gender.

Such people live their childhood in frustration, knowing something is wrong but being unaware of the specific issue. Finally, due to logical reasoning or reading of someone else's experiences, the child or young adult awakens to the truth. The inner conflict is unrelated to behavior (lifestyle) or preference of sexual partner; rather, it is a conflict involving one's core identity. Medical terms such as gender dysphoria or gender identity disorder (GID) are sometimes used to describe the condition, but it is perhaps best known by the popular name, transsexualism.

Childhood Experiences of Transsexual People

Most adult transsexual people can remember feeling different from their same-sex peers. Before they knew the physical differences between boys and girls, and long before they began to experience feelings of sexual attraction, they knew they belonged in the opposite camp. Such knowledge can be terrifying in an environment where conformity is demanded and diversity is rejected.

Gender-variant behavior in children is not uncommon. Tomboy behavior in girls is usually tolerated better than effeminate behavior in boys. Many boys who play with dolls and other girls' toys grow up to become heterosexual men; some grow up to be gay men; and a few never consider themselves men at all, but they may fear for their safety if other children regard them as sissies. Childhood conformity and peer pressure are the first obstacles transgender people encounter in their journey toward transition.

Factors that Discourage Transition

Even stronger than a child's peer relationships are the bonds of family. Parents, brothers, and sisters are the most influential people in any child's life. In contrast to the previously held views of some mental health professionals, who theorized that transsexualism results from family dysfunction in early childhood, many transsexual people report normal and loving family relationships. Even so, it is rare for a transsexual child to feel comfortable discussing gender feelings with family. Perhaps the child loves his or her parents and fears hurting them with so startling a revelation; or perhaps the fear of punishment creates a determination to hide the truth. Such fears are not without merit – it is true that many parents are ill prepared to deal with such a revelation by a child.

Psychological stress is a natural result of the child's practice of denial, which many continue well into adulthood. Some children are so affected by the stress that their school progress is impaired. Others find temporary escape in retreat into books and studies. Music or other hobbies may serve as a welcome distraction from their worries. Some even practice denial so strongly that they seek the other extreme; for example, a boy may strive to excel at sports or to engage in high-risk behavior, placing himself in physical danger, to deny his feeling of being a girl. Indeed, some teens and young adults enlist in the military services in an effort to cure such feelings.

Even in childhood, religion may play a large role in the stress a transsexual child faces. Most religions teach absolute truths, and do not tolerate departures from those absolutes. The child is taught that his or her feelings are sinful or evil but that they can be overcome through the faith process. Most children are eager to believe such doctrine, hoping that they can be made normal so their parents and peers will accept them. When the religious experience fails to change a child's gender identity, the child is left with feelings of failure and self-doubt. A significant number of transsexual adults report so much despair as children or teenagers that they considered suicide.

The onset of puberty is a very stressful time for transsexual children. Their dreams of becoming normal members of their desired gender are severely challenged as female-to-male (FTM) people develop breasts and menstrual activity and male-to-female (MTF) youth experience growing to tall stature, beard growth, and a deepening voice. These physical changes require years of treatment through surgery, hormones, and electrolysis to correct – if they can be corrected at all.

The ultimate denial of an individuals' transsexualism focuses on their attempt to live a normal life through marriage and parenthood. In years past, transsexual young adults lacked the information and resources available today. Believing themselves doomed to a life of quiet unhappiness, they pinned their last hopes on marriage and family to cure them of their dysphoria. However, within a few years, they realized this too could not cure their sense of identity. Unfortunately, at this point other people have entered their lives in intimate roles, and any ultimate disclosure and transition will be disruptive to the entire family. Because of this, some people make the decision to delay transition until their children are older; this is the reason we see so many people begin transition after age 40. A decision to delay carries its own dangers; a child who reaches puberty before learning of a parent's transsexualism usually will react more negatively than a child who is made aware at a younger age.

The Transition Process

Eventually the transsexual person reaches a point of understanding – this is the way my life is going to be. Denial, family pressure, religious fervor, and even marriage have all failed to change the person's gender identity. The incredible pressure can be relieved only through acceptance and transition. The process of disclosure and beginning transition is sometimes called coming out.

This is perhaps the time of the most intense stress a transsexual person experiences. Every aspect of life is at risk: family, friends, status, job, and finances. The importance of advance planning becomes crucial. People who felt they would be warmly accepted may be devastated by the rejection they experience. Job loss, although not universal, is still very common. Income is gone, at a time when it is needed more than ever. The expenses of transition can be enormous; the psychological counseling alone is more than some people can afford, and the various surgical procedures may cost tens of thousands of dollars and very few are covered by insurance. Efforts to find a new job during transition, in the new gender role, are made more difficult because a lifetime of documentation – birth certificate, Social Security records, passport, diplomas, credit ratings – are all on record under the old name. Furthermore, a person in early transition still has behavior patterns from a lifetime of socialization in the birth gender. Such behavior may be overcome quickly for some people as they find the freedom to behave naturally as a member of their desired gender, but others struggle for years to shed male mannerisms. When added to a physical appearance that still reflects the birth gender, the new mannerisms make it difficult for some to blend into society without attracting unwelcome attention.

For some, the stresses of transition prove too great. As a result of economic demand, family pressure, or the hope of a religious cure, some people abandon transition and return to their original gender role. For most, this return is temporary, and the transition process resumes once they are better prepared to cope with the stress.

Hormonal and Surgical Changes

During and after transition, the effects of cross-gender hormone therapy produce results that often are dramatic. FTM people experience a deepening of the voice and the growth of facial hair. The most obvious early effect of hormones for MTF people is breast enlargement. These effects are obvious to friends, family, and coworkers. At this point, transsexual people may feel a sense of relief because they no longer have to hide their identity or their plan to transition.

Of course, hormones do not produce a complete transformation to the new gender. Other physical differences, including skull and jaw structure, cannot be reversed by hormone treatment. Surgical procedures to eliminate prominent male jaw and chin shape, as well as the exaggerated prominence of the brows, can be of great help in allowing MTF people live normal lives without constant exposure. FTM people often seek mastectomy to make life less stressful.

Ultimately, the transition process culminates in sex reassignment surgery (SRS). In this operation, the surgeon removes some of the old sex organs and creates an appearance as close as possible to the desired gender. Most surgeons who perform SRS follow rather strict guidelines of patient selection to minimize the risk of postoperative regret.

Health Effects of Stress for Transsexual People

A number of physical and psychological effects may result from a transsexual person's decisions related to transition. Many effects occur prior to coming out. The stress of keeping hidden a life-changing truth may be a factor in all types of stress-related illness, including high blood pressure, migraine or other headaches, peptic ulcer disease, and inflammatory bowel disease.

The psychological impact of stress often manifests as chronic anxiety or depression. The degree of psychological distress may be so severe that the person cannot function normally in a work or school

environment. Anxiolytic medications, especially benzodiazepines, are often prescribed; even more commonly given are antidepressants. Some degree of depression is extremely common in transsexual people who have experienced rejection or loss.

Suicide, or suicidal gesture, is a continuing concern for many transsexual people. There are three periods when transsexuals are at greatest risk for suicide. The first occurs early in life, when the young person experiences despair from being unlike his or her peers; the stress of keeping such knowledge undisclosed is magnified by the failure of attempts to eliminate the dysphoria. The second period of high suicide risk occurs just prior to the person's acceptance of his or her transsexualism and decision to proceed with transition and is associated with the last attempt to avoid disclosure and to have, to outward appearances, a normal life. As already mentioned, acceptance and transition may relieve this stress and the associated depression.

The third period of high suicide risk occurs after the completion of transition. The person may have high expectations of success in the chosen gender, and some expectations are more likely to be realized than others. Most people find acceptance by society and may continue their careers or begin new ones. Some, however, lose their jobs and are unable to find employment sufficient to pay expenses. Family and friends may continue to reject the transsexual to the point of shunning the person. Perhaps most devastating, the transsexual may be unable to find a life partner whose love and acceptance is unconditional.

The psychological counseling included in a well-planned transition process should address these expectations so that transsexuals can consider how to deal with failure to meet them. Some people, unfortunately, view the counseling process as simply another obstacle to be overcome in order to obtain their desired surgery. The lack of attention to good counseling during transition may increase the likelihood of regret, depression, and even suicide in the years that follow.

Religion and the Transsexual Person

It is accurate to generalize that people whose religious heritage is liberal, or who have no religious affiliation, experience less stress with transition than people who come from a conservative religious background. This generalization can apply to conservative Christianity, and also to Judaism and Islam. Some churches, particularly the Roman Catholic church, have made specific doctrinal statements describing transsexualism as unacceptable. Other Protestant churches may include transsexualism with homosexuality and issue a general condemnation of all gender and sexual diversity.

People who have been raised in such denominations experience a high degree of stress. Their religion tells them it is a sin to change from one gender to the other, and it threatens them with eternal punishment if they disobey. The church maintains that a person can, through faith, turn away from transsexualism and live a normal life. For many, the church is their major support network; the thought of being cast out is unbearable. Yet their life experience tells them otherwise – they have already experienced failure of their faith to bring about this change to normalcy.

These people are forced to choose between two difficult paths. Either they must continue to live in denial, experiencing the stress of secrecy and eventually reaching the point of having to face the truth, or they must face the truth early, risking probable rejection and becoming disillusioned with their faith. Many such people do find their way to more liberal, accepting churches, but others abandon their faith altogether and may become quite hostile to any form of religion.

Violence and Fear

It is not coincidence that transsexual people are more likely to be the victims of violent crime. Several factors explain this increased risk. Young, aggressive males, the group most likely to show violent behavior, have a high incidence of homophobia. If a homophobic male encounters someone who is visibly gender variant, he may feel his own masculinity is threatened; he may even feel the need to prove his manhood by attacking the queer.

In addition, when a person is not initially perceived as transsexual but is discovered to be so later, there may be anger because of the perception of having been deceived. Well-publicized cases of people such as Gwen Araujo and Brandon Teena, both of whom were murdered under these circumstances, illustrate the dangers of antitranssexual violence.

Such is the reality faced by transsexual people in our society. Most people take a healthy approach to such risks by avoiding situations in which the danger is apparent. Others may avoid all social occasions and relationships; their excessive fear of violent reactions prevents their having a normal social life.

Intimacy and the Transsexual Person

Many transsexual people find themselves alone at the beginning of transition. Either they have not been in a relationship or their spouse has refused to be a part of

their future plans. These people can benefit from the transition counseling process if they are willing to take the time to explore their future intimacy needs. Such exploration can be stressful in itself, however. People in transition learn that gender identity is independent of sexual orientation. An MTF transsexual person, for example, who was attracted to women prior to transition may ultimately seek intimacy with a man or with a woman. The problem comes when the desired partner has to deal with the reality of the transsexual person's past. A relationship with a man is often jeopardized by the man's feelings of homophobia and inability to get past the idea that his partner is not a real woman. Less likely, but also possible, is the failure of a lesbian relationship due to the female partner's similar feelings. It is possible to find love posttransition, but the increased difficulty of doing so is a great source of stress.

A different dynamic occurs when the spouse of a transsexual person chooses to remain in the relationship posttransition. What was initially a heterosexual relationship now is seen by the world as a homosexual one, with all the associated societal stress such couples encounter. The continued relationship becomes quite a strain on the spouse; for example, the wife of a MTF transsexual soon realizes that she cannot change her sexual orientation to lesbian, just as a gay person cannot decide to become straight. Couples who stay together after transition are the rare exception, and in general the two people involved are very mature and are strong and secure in their commitment to one another despite the stress of daily confrontation.

Aging and Death and the Transsexual Person

The subject of aging has only recently been discussed in depth among transsexual people because the number of people completing transition remained small until the 1980s. Now some people from that era are entering their 60s and 70s, and aging and health issues have begun to assume a prominence in their priorities and present unique stress situations.

Many elderly people count on their children or grandchildren for physical support, to attend to their medical needs, and to take them for doctor visits or hospital emergencies. Transsexual people may lack that support because of family rejection. They face the stress of finding other support among friends, and sometimes a network of elder transsexuals may form for mutual support. The idea of loss of independence is stressful for everyone, but especially so for people who may wear breast binders or hair-replacement systems.

The medical needs of the aging transsexual people are often unique. Physicians called to see a transsexual patient in an emergency may not possess the knowledge or the desire to provide appropriate treatment for these unique needs. It may be difficult to find a doctor willing to continue hormone therapy in an older individual or to screen for diseases associated with the person's birth gender.

A case was reported of an FTM person who developed ovarian cancer. He was rejected for treatment by several physicians, who could not accept a person who presented as a male but had ovaries. The patient eventually died from his disease, which had gone untreated unnecessarily. It is also possible (unlikely, but reported) for an MTF person to develop cancer of the prostate, and it is very rare for such people to submit to the usual screening blood tests.

End of life issues need to be addressed in transsexual people with special concern. It is very important that transsexual people have a partner or sympathetic friend to serve as health-care power of attorney. Likewise, it is important for them to have a will specifying how they wish to be buried or cremated. This will avoid the situation in which family members disregard the individual's stated wishes and bury him or her as the birth gender, even many years posttransition.

Conclusion

Transsexualism, more common than once thought, affects thousands of people who are subject to increased stress due to rejection or confrontation from family, friends, employers, religious institutions, and even total strangers. Coping with these stresses on a daily basis may make the transsexual person mature and strong, or it may be devastating. Understanding and acceptance may reduce the constant stress and enable the transsexual person to make valuable contributions to society.

Further Reading

Boylan, J. F. (2003). *She's not there: a life in two genders*. New York: Broadway Books.

Brown, M. L. and Rounsley, C. A. (1996). *True selves: understanding transsexualism for families, friends, coworkers, and helping professionals*. San Francisco, CA: Jossey-Bass.

Cromwell, J. (1999). *Transmen and FTMs: identities, bodies, genders, and sexualities*. Urbana, IL: University of Illinois Press.

Ettner, R. (1996). *Confessions of a gender defender: a psychologist's reflections on life among the transgendered*. Chicago, IL: Spectrum Press.

Ettner, R. (1999). *Gender loving care: a guide to counseling gender-variant clients*. New York: W. W. Norton.
Green, J. (2004). *Becoming a visible man*. Nashville, TN: Vanderbilt University Press.
Israel, G. E. and Tarver, D. E., II (1997). *Transgender care: recommended guidelines, practical information and personal accounts*. Philadelphia, PA: Temple University Press.
Kaiser Permanente National Diversity Council (2004). *A provider's handbook on culturally competent care: lesbian, gay, bisexual, and transgender population* (2nd edn.). Oakland, CA: Kaiser Permanente National Diversity Council.
Lev, A. I. (2004). *Transgender emergence: therapeutic guidelines for working with gender-variant people and their families*. New York: Haworth Press.
Stryker, S. and Whittle, S. (2006). *The transgender studies reader*. New York: Routledge.

Trauma and Memory

B A van der Kolk
Boston University School of Medicine, Boston, MA USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 620–622, © 2000, Elsevier Inc.

Trauma and Dissociation
The Nature of Traumatic Memories
Amnesia and the Return of Dissociated Memories
The Neurobiology of Traumatic Memory

The issue of trauma and memory is central to the concept of posttraumatic stress disorder (PTSD). The *Diagnostic and Statistical Manual of Mental Disorders* (4th edn, DSM-IV) defines PTSD as being accompanied by extremes of remembering and forgetting. Because traumatic memories take on their extreme character only after a person has been exposed to overwhelming stress, laboratory studies using conventional stimuli seem to be of little guidance in clarifying what happens when people are traumatized. Because trauma cannot be simulated in the laboratory, the exploration of the relationship between trauma and memory depends on (1) collecting retrospective reports from traumatized individuals, (2) post hoc observations of people with trauma histories, or (3) provoking traumatic memories and flashbacks in the laboratory.

Trauma and Dissociation

Christianson and others have described how, when people feel threatened, they tend to experience a significant narrowing of their consciousness while remaining focused on the central details of their experience. This increasing narrowing of consciousness is a refection of the dissociation of traumatized individuals. Dissociation at the moment of the trauma has been found to be the single most important predictor of the subsequent development of PTSD. Dissociation has been described as a concomitant of trauma during World War I and II and after accidents, rapes, and numerous other traumatic experiences. Although dissociation may have some adaptive value, it also seems to lead to a disintegration of the capacity to integrate the different perceptual elements of the traumatic experience. Extreme physiological arousal interferes with proper information processing and with the storage of information into narrative (explicit) memory. When people are overwhelmed, they may suffer from speechless terror in which words fail to describe what has happened. However, even though traumatized individuals may be unable to make a coherent narrative of the traumatic events, there may be no interference with implicit memory; they may know the emotional valence of a stimulus and be aware of the associated perceptions without being able to articulate the reasons for feeling or behaving in a particular way. Traumatic memories are split off from (dissociated) consciousness and are stored as sensory perceptions, obsessional ideas, or behavioral reenactments.

The Nature of Traumatic Memories

Under ordinary conditions, memory always is an active and constructive process. Whereas familiar and expectable experiences are generally easily integrated and memories of ordinary events are bleached in clarity over time, some aspects of traumatic events appear to get fixed in the mind, unaltered by the passage of time or by the intervention of subsequent

experience. However, the emotional valence of experience affects the accuracy of memory; reports of personally highly significant events are often unusually accurate and remain fairly stable over time. Evolution favors the consolidation of relevant information. Since late in the nineteenth century, clinicians have noticed over and over again that the memories of traumatized patients seem to consist predominantly of emotional and perceptual elements, whereas there appears to be an impairment of declarative memory. A series of studies on the memories of traumatized populations have demonstrated that, regardless of the nature of the trauma, traumatized individuals initially report having decreased verbal recall, whereas they are prone to reexperience the traumatic events in the form of visual images, kinesthetic sensations, sounds, smells, and extreme affective states. Whether we study people reporting childhood trauma, car accidents, rape, or other assaults or people who wake up in the middle of a surgical procedure, this pattern of recollection is consistently reported by victims with PTSD. This pattern of memory recollection suggests that traumatic memories are processed differently in the brain than memories of day-to-day experiences.

It is generally assumed that memories can easily be distorted and that verbal recall changes over time. Even so-called flashbulb memories tend to be modified with the passage of time. However, traumatized individuals tend to report that the content of nightmares and flashbacks remain the same. The author was able to show that memories of traumatic experiences change after effective treatment of PTSD but do not change if the PTSD remains untreated. There have been no reports in the scientific literature showing that the sensory and perceptual elements of traumatic memories change in quality with the passage of time; nor has the reverse been reported – that these elements of traumatic memories always remain unaltered.

Amnesia and the Return of Dissociated Memories

Amnesia following adult trauma has been very well documented following World War I and II, in accident victims, in concentration camp survivors, and in victims of torture. In recent years, the occurrence of amnesia following childhood maltreatment has been called into question. This controversial issue has given rise to numerous studies. At this time, there have been 39 published reports in peer-reviewed journals that have documented that anywhere between 20 and 80% of people with sexual abuse histories suffer from amnesia for the event at some time during their lives; every published study of this population has found some percentage of subjects who suffered from amnesia. Precipitants for the return of these previously forgotten memories tend to be events or sensations that are reminders of the original trauma. Events such as witnessing someone else going through trauma similar to one's own, hearing about the similar trauma, seeing a movie, or being in the company of someone who reminds a person of him- or herself at the same age at which the abuse occurred tend to be the most common precipitants for the return of these memories. A well-known study by Linda Williams of a group of women with documented sexual abuse histories found that the memories of the women who had a period of amnesia and who later recalled their trauma were at least as accurate as the memories of the women who had always recalled their abuse.

The Neurobiology of Traumatic Memory

Based on animal research, it has been widely assumed that the massive secretion of neurohormones at the time of the trauma plays a critical role in the ways that traumatic memories are consolidated into long-term memory. The strength of hormonal stimulation at the time a particular event occurred determines how strongly the memory is laid down. This capacity helps organisms evaluate the importance of sensory input better and causes emotionally significant material to be accessed more easily. The problem in PTSD is that the capacity to access memories related to the trauma becomes too easy; the trauma is recalled whenever a person is exposed to a sufficient number of matching sensory stimuli, regardless of their current relevance.

Traumatic Memories Are State-Dependent

In people with PTSD, sensations reminiscent of the trauma, including states of high physiological arousal, selectively promote the retrieval of traumatic memories or precipitate behaviors associated with the original trauma. Classically, a Vietnam veteran may have a flashback during a Fourth of July picnic and behave as if he were back in Vietnam, but he might not respond to fireworks on New Year's Eve because only the sound, but not temperature, matches the sensations that accompanied the traumatic events. Biological research has shown that drugs that precipitate autonomic arousal, such as lactate or yohimbine, tend to precipitate panic attacks and flashbacks in patients with PTSD. These can be considered instances of state-dependent memory retrieval.

Functional and Neuroanatomical Correlates of Posttraumatic Stress Disorder: Implications for Understanding Traumatic Memories

A range of both neuroanatomical and functional changes have been documented in individuals with PTSD. Several studies have shown decreased hippocampal volume, as well as decreased volume of the corpus callosum in patients with abuse histories. Functional neuroimaging has shown that people with PTSD have an increased activation of the right amygdala when exposed to a traumatic stimulus, which is accompanied by a decreased activation of Broca's area. These findings suggest that the difficulty that many PTSD patients have in putting their memories into words reflect in abnormalities in brain activation. Although it is unclear at the present time what the precise clinical implications of these findings are, it is tempting to speculate that a decreased hippocampal volume may be responsible for the fragmented nature of traumatic memories. Altered cerebral lateralization, with increased right hemispheric activation, may be relevant for understanding dissociative phenomena.

See Also the Following Articles

Amnesia; Concentration Camp Survivors; Dissociation; Memory and Stress; Posttraumatic Stress Disorder, Neurobiology of; Torture.

Further Reading

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th edn.). Washington, DC: American Psychiatric Association.

Appelbaum, P. (1998). *Trauma and memory: clinical and legal controversies*. Oxford University Press.

Bremner, D. and Marmar, C. (1998). *Trauma, memory, and dissociation*. Progress in Psychiatry, no. 54. Washington, DC: APA Press.

Brown, D. P., Scheflin, A. W. and Corydon Hammond, D. (1998). *Memory, trauma treatment, and the law*. New York: W. W. Norton.

Rauch, S., van der Kolk, B. A., Fisler, et al. (1996). A symptom provocation study using positron emission tomography and script driven imagery. *Archives of General Psychiatry* **53**, 380–387.

van der Kolk, B. A., McFarlane, A. C. and Weisaeth, L. (eds.) (1996). *Traumatic stress: the effects of overwhelming experience on mind, body and society*. New York: Guilford Press.

van der Kolk, B. A., Burbridge, J. A. and Suzuki, J. (1997). The psychobiology of traumatic memory: clinical implications of neuroimaging studies. *Annals of the New York Academy of Sciences* **821**, 99–113.

Trauma Group Therapy

J Dwyer and L G Martin
VA Greater Los Angeles Healthcare System, Los Angeles, CA, USA

© 2007 Elsevier Inc. All rights reserved.

History
Characteristics of Trauma Groups
Definition of Group Therapy
Group Process
Trauma Group Structure

Glossary

Arousal	A state of activated emotional activity.
Catharsis	Psychotherapy that encourages the discharge of pent up, socially unacceptable affects; the discharge of pent up emotions to result in the alleviation of symptoms or permanent relief of the condition.
Cognitive restructuring	The identification and modification of maladaptive thought processes.
Exposure	The disclosure of something private or secret; the act or instance of revealing or unmasking, for example, traumatic events.
Process	A systematic series of actions directed to some end (e.g., the uncovering of repressed emotions and memories).
Rap group	Peer group discussion.
Talking cure	The verbal interaction between therapist and patient (or peers) regarding the goal of uncovering hidden trauma; coined by Sigmund Freud.

History

Before presenting the elements of the trauma group, it is essential to understand the efforts to legitimize the trauma syndrome itself. Society has moved into a period when people are seeking the human suste-

nance that has been lost in the development of technology, the emergence of the megalopolis, and the increased complexity and depersonalization of societal structures. Individuals and communities are too often witnessing and experiencing the traumatic effects of the big hurts in life. Generalized angst in our societies prompts a search for a sense of community with intimate and mutually supportive human ties, a search for community that enhances the lives of individuals through a common purpose. Toward this end, more people are forging their own microcosms of community in the form of therapy or support groups. These come in many forms: action groups, which support collective or individual impact on societal institutions; interpersonal exploration groups, in which individuals seek to determine who they really are versus how they appear to others; and other groups pursuing strongly held common interests such as learning, experiencing or performing in a world of ideas and skills.

In the late nineteenth century, the disorder known as hysteria became the focus of a major inquiry. French neurologist, Jean-Martin Charcot studied the symptoms of hysteria, which resembled neurological damage: motor paralysis, sensory loss, convulsions, and amnesia. By 1880, he had demonstrated that these symptoms were psychological and could be artificially induced and relieved using hypnosis. Charcot's followers pursued the cause of hysteria. Rivalry in this pursuit was particularly intense between Janet and Freud. By the mid-1890s, Janet in France and Freud (with his collaborator Joseph Breur) in Vienna had independently arrived at strikingly similar formulations – hysteria was a condition caused by psychological trauma. While conducting explorations toward a cure, they found it necessary to speak to their patients. They found that unbearable emotional reactions to traumatic events produced an altered state of consciousness, which, in turn, yielded the symptoms of hysteria. Janet described these patients as being governed by subconscious fixed ideas – the memories of traumatic events. Breur and Freud, in an immortal summation, wrote that hysterics suffer from reminiscences.

By 1908, these investigators had also discovered that hysterical symptoms could be alleviated when the traumatic memories and intense feelings that accompanied them were recovered and put into words. Janet called the technique psychological analysis; Breur and Freud dubbed it the talking cure. Freud later recanted the link between hysteria and trauma, and the entire line of inquiry fell into neglect. The study of trauma came to a halt.

Reemergence of Trauma Investigation

The reality of psychological trauma was forced into the public consciousness again by the catastrophe of World War I, although the primary mental health conclusion at that time was that soldiers who developed a traumatic neurosis were at best constitutionally inferior human beings and at worst cowards and malingerers. Within a few years after World War I, medical interest in the subject faded again. In 1922, a young American psychiatrist, Abram Kardiner, took up the case of traumatized soldiers. He developed the clinical outline of the trauma syndrome as it is understood today.

With the casualties of World War II came a revival of medical interest in combat neurosis. Various studies led to the conclusion that the strongest protection against this overwhelming terror was the degree of relatedness between the soldier, his immediate fighting unit, and their leader. Their microcosmic community recreated treatment strategies, evolved during World War II, that were designed to minimize the separation of the afflicted soldier and his comrades. In the quest for a fast and effective method of treatment, hypnosis was used, as it had been in earlier work on hysteria. Under hypnosis, the talking cure was used with a focus on recovery and cathartic reexperiencing of traumatic events and their attendant feelings of terror, rage, and grief.

Systematic large-scale studies of the long-term effects of combat were not undertaken until after the Vietnam War. The use of groups to treat trauma was pioneered with Vietnam veterans. By the mid-1970s, hundreds of informal rap groups had been organized.

Characteristics of Trauma Groups

From their beginning, trauma rap groups were different and innovative. They had three guiding principles: (1) The groups were formed around an affinity, or the coming together of people who shared an historical perspective or personal experience; (2) the task of the facilitator was to be empathic and manifest a presence at all times of full engagement, openness, and interaction and not to serve merely as a sounding board; and (3) the veterans had to assume responsibility for self-generation, that is, to initiate their own process and conduct groups largely on their own terms.

Early proponents of trauma treatment in the group model illuminated six characteristics that are shared by veterans active in the groups.

1. They have a pervasive guilt from which they cannot escape. Its nexus resides in having lived

when others from both sides of battle have been killed or severely maimed. The guilt leads to a pathological desire to atone for the sin of having survived. To this end, they become involved in reckless forms of self-punishment such as substance abuse or one-car accidents that are just a shade less determined than actual suicide.

2. They have a feeling of being scapegoats, having been used and betrayed.

3. They have a generalized rage at society and at the individuals they believe have duped and/or manipulated them.

4. They have undergone combat brutalization, a method of military training aimed at creating the most effective soldiers by numbing the emotional processes that could interfere when the need to fight or kill arises. This learned detachment leads to the fifth characteristic.

5. They are alienated from their own humanity and withdrawn from others. This alienation inevitably leads to the sixth and most devastating characteristic of all.

6. They have a profound inability to love, accept affection, or to be intimate with others.

These characteristics predominated then, and they prevail today as the symptoms we most often see presented in treatment for trauma.

Definition of Group Therapy

Group therapy is psychotherapy practiced ideally in a group of 8–10 members/participants. The group itself becomes an important entity in the treatment of its individuals. The group collective assists its members by guiding, constructively criticizing, and offering therapeutic suggestions and support. It is the application of psychotherapeutic techniques within a small group of individuals who share similar experiences and have difficulty in responding to them. Each group has one or two leaders or facilitators whose role it is to guide and direct the group discussion or processing of problems in an attempt to promote personal growth and favorable behavioral change.

Types of Therapy Groups

There are different kinds of therapy groups. They may deal with a variety of issues including interpersonal conflicts, addictions, or trauma. The Department of Veterans Affairs National Center for Posttraumatic Stress Disorder (PTSD) has written that the best option for most PTSD patients is group therapy. In the trauma group, patients discuss with others their traumatic memories, symptoms of PTSD, and functional deficits. This approach has been most successful with combat veterans, rape and/or incest victims, and natural disaster survivors. Indeed, the group model has been so effective that it lies at the core of the ever-growing self-help movement. The trauma group promotes discovery of, insight into, and eventually integration of the traumatic event.

Group Process

Trauma groups employ two important aspects of trauma recovery: the talking cure and the healing benefits of modeling supportive interactions with others in small-group cohesion. Survivor groups typically comprise 8–10 people meeting regularly to provide support and encouragement to one another. They accomplish a number of steps in healing by confronting alienation and estrangement. Survivors begin to realize that they are not alone in their reactions. Their feelings are normalized, and they begin to feel that they are part of something rather than being different or separate. The group offers a sense of community and a feeling of security.

Cognitive Restructuring

An important function of the trauma group is to destigmatize the traumatic experience. It focuses on exposure, which is the repeated disclosure of the event(s), and guides patients to a more realistic view of themselves and their world (cognitive restructuring). In addition, the trauma group combines the strengths of its members by having them each share coping skills and ideas that have benefited them. Most important, the group creates a milieu in which members can learn to trust and establish an ability to enter into relationships again. The group provides a laboratory for trusting others and, in time, patients can generalize this trust and begin to form healthy relationships in the outside world. The group model has been shown to be inappropriate and not helpful with patients who are suicidal, homicidal, violent, actively abusing drugs, or self-mutilating.

Trauma Group Structure

The foundation of the trauma group rests on exposure, cognitive restructuring, and the use of the traumatic narrative. It allows for reexperiencing trauma in manageable doses. The methods used in each phase share the common feature of helping anxious patients confront their most fearsome recall with the goal of reducing irrational fear (anxiety). Exposure has been the standard of treatment for anxiety disorders for several decades. The trauma group, in phases, employs both exposure and cognitive therapy to extinguish the prevailing pathological fear

structure in the survivor's memory. The effective group process offers the patient corrective trauma-related information. This, in turn, provides an opportunity for patients to begin the process of desensitization, or adaptation to the trauma.

Repeated reliving of the trauma reduces it to a specific occurrence rather than a generalized representation of the world as a universally dangerous place and reduces patients' self-image of incompetence. The trauma group is structured to merge exposure with cognitive therapy; this merger yields the desired therapeutic effect – the integration of dissociated trauma material and newly acquired corrective information.

Phases of the Trauma Group

Phase I The first phase of the trauma group addresses the issues of trust and personal safety. Patients are given assurance that group content is absolutely confidential and that no violence will occur during group sessions. Participants learn quickly that they are in control and that the model involves small steps toward emotional passages. "Managing small doses of awareness" becomes the mantra of the group. Patients know that they are in control, an essential element derived directly from the rap group model. During this phase, patients build trust with the facilitator(s) and other group members. They learn that the group medium fosters support and mutual healing, and the trust emerges slowly, with each emotional victory. Emphasis is on the present – the here and now. The importance of self-care, stress management, and anger management is highlighted. The process confronts the Lone Ranger syndrome and encourages each member to become part of the cohesive unit. The beginning steps of trust and bonding begin to surface; healthy and supportive relationships are born and nurtured. The patients demonstrate emotional growth, indicated by patience, courage, and the desire to confront their fears. They have constructed a healthy emotional platform from which to begin the arduous task of thawing freeze-dried traumatic memories.

Phase II The second phase of the trauma group is the narrative/trauma phase. Survivors have stories that must be revealed to themselves and to others. The fundamental human need to share and connect (Maslow) is enacted in phase II. It becomes about exposing and facing the hidden horror. Initially, narratives can be fragmented and distorted. Refining them in a controlled, supported environment reveals the conflicts between the memory and the actual event(s). The narrative helps to thaw out the frozen memory. It offers an important way to desensitize patients to the enormous emotional overload. By repeated telling of the story and significant group feedback, a linear narrative emerges and patients begin to see the event more clearly and can begin to integrate the frozen memory into their broader life picture.

Phase III The last phase of the trauma group is the reconnection phase, or Life 101. A desire and ability to return to a productive, healthy, and satisfying life style returns. The symptoms of PTSD have become manageable. Patients can bear the feelings associated with the traumatic memories because they are now in a position of authority over them. The memory, with all of its dreaded darkness has become a coherent story linked to an appropriate affect. Patients, unlike the carpenter with a hammer, no longer see everything as a nail. Long-held perceptions regarding self and the world have been altered. Often, the survivors can rid themselves of the hero–villain paradigm of life. They can once again embrace a humanity that has been out of reach for so long. Patients begin to accept that life is imperfect and that in life both love and horror can coexist. Patients move from alienation from themselves to being the narrator of their own lives. Reconnecting means repairing the covenant of trust with themselves, which was broken at the time of trauma. It means being active, engaged, and prosocial. It means serving others and striving to make the world a better place. Life can become meaningful and fulfilled with a balance of joy, laughter, and sometimes tears. Most important, survivors can once again give and receive love. There is life after trauma.

Further Reading

Briere, J. (2004). *Psychological assessment of adult posttraumatic states*. Washington, DC: American Psychological Association.

Foa, E., Keane, T. and Friedman, M. (eds.) (1998). *Effective treatments for PTSD*. New York: Guilford Press.

Follette, V., Ruzek, J. and Abueg, F. (eds.) (1998). *Cognitive behavioral therapies for trauma*. New York: Guilford Press.

Herman, J. (1992). *Trauma and recovery*. New York: Basic Books.

Kardiner, A. and Spiegel, H. (1947). *War, stress and neurotic illness*. New York: Hoeber.

Schiraldi, G. (2000). *The PTSD source book*. Los Angeles: Lowell House.

Shatan, C. (1973). The grief of soldiers: Vietnam combat veteran's self-help movement. *American Journal of Orthopsychiatry* **43**, 640–653.

Shay, J. (1994). *Achilles in Vietnam*. New York: Simon & Schuster.

Williams, M. B. and Sommer, J. (eds.) (1994). *Handbook of posttraumatic therapy*. Westport, CT: Greenwood Press.

Traumatic Stress and Posttraumatic Stress Disorder, the Israeli Experience

E Klein
Faculty of Medicine Technion, Haifa, Israel
J Zohar
Tel Aviv University, Tel Aviv, Israel

© 2007 Elsevier Inc. All rights reserved.

Introduction
Posttraumatic Reactions in Holocaust Survivors
Combat Stress Reaction and Posttraumatic Stress Disorder in War Veterans
The Impact of Terror on Traumatic Stress Responses in the Israeli Population
Neurobiology and Psychophysiology of Posttraumatic Stress Disorder
Animal Model for Posttraumatic Stress Disorder
Conclusion

Glossary

Acute stress disorder	An anxiety state lasting up to 4 weeks following a traumatic event.
Acute stress reaction	An acute and time-limited (up to 48 h) anxiety state following a traumatic event.
Animal models	Experiments in which animals are used to mimic and study human diseases.
Combat stress reaction	An acute stress reaction that develops on the battlefield.
Gulf War	The 1991 war in which the United States and the allied countries fought against Iraq following its invasion of Kuwait.
Intifada	Recent Palestinian campaigns directed at ending the Israeli military occupation (from انتفاضة *intifāḍah*, shaking off; Arabic for uprising).
Posttraumatic stress disorder (PTSD)	A chronic disorder that develops in survivors of traumatic events.
Stressful life event	An event that might cause emotional distress.
Traumatic event	A life-threatening event.
Yom Kippur War	The 1973 war in which Egypt and Syria attacked Israel.

Introduction

Emotional trauma and stressful life events have been implicated in the development of psychiatric disorders, particularly anxiety and depression. Traumatic events differ from stressful life events in that they involve a direct threat to life or physical/psychological integrity. Posttraumatic responses can be diverse, but the most typical form of psychiatric morbidity resulting from exposure to a traumatic event is posttraumatic stress disorder (PTSD). It is estimated that at least 50% of the people in Western countries are exposed to traumatic events during their lifetime. Because PTSD occurs in approximately 10–20% of the individuals who are exposed to a traumatic event (the majority eventually adapt), the estimated life prevalence is 10% of 50% (i.e., 5%). Indeed, epidemiological studies show that 1–3% of the adult population in the United States suffer from PTSD. The finding that only a fraction of those exposed to traumatic events develop PTSD suggests that the traumatic event is necessary but not sufficient for its development and that additional factors are important in determining the individual vulnerability of trauma survivors to developing PTSD.

The study of human responses to trauma has been intimately associated with the history of wars in the twentieth century. There have been anecdotal accounts of posttraumatic responses in soldiers in the American Civil War and other wars in the nineteenth century, but World War I (1914–1918) is clearly the cradle of the modern study of trauma and its consequences.

Only in the last 2 decades has trauma research evolved to focus on civilian populations that have been exposed to violence, sexual abuse, terror activity, traffic accidents, and health conditions (e.g., heart attacks), which have all been shown to cause PTSD and are responsible for the high prevalence of the disorder in civilians.

Israel has become, since its establishment in 1948, a live laboratory for the research of trauma and PTSD. The horrors of the Holocaust and its emotional scars in the survivors, followed by seven wars and waves of terror, have unfortunately served as an experimental field for the study of trauma and its consequences. This article reviews key findings from studies done in Israel on trauma and PTSD during the last 2 decades and highlights their contribution to the existing body of knowledge in this field.

Posttraumatic Reactions in Holocaust Survivors

The Holocaust is an unparalleled event in human history that has left enduring emotional scars in all

its survivors. Most Holocaust survivors were able to readjust surprisingly well and restore their lives, whereas others were not able to overcome the dreadful events, which continued to cast a shadow on their entire life, making it impossible for them to regain security and emotional well-being.

Studies of Holocaust survivors were done mainly in the 1960s and 1970s, before the era of modern trauma research and the conceptualization of the definitions currently used in the field. This, and the fact that many of the survivors have already died and those who are still alive are in their 80s, explains the relative paucity of updated data on posttraumatic responses in Holocaust survivors.

Combat Stress Reaction and Posttraumatic Stress Disorder in War Veterans

The Yom Kippur War and Lebanon Wars

In contrast to the earlier wars in the short history of the state of Israel (1949, 1956, 1967), which were not studied systematically in the context of traumatic stress, some of the soldiers with combat stress reaction (CSR) during the Yom Kippur (1973) and Lebanon (1981) wars were identified early on, received treatment, and were followed over time, so that data could be collected about the short- and long-term consequences of exposure to traumatic events during wartime. The Lebanon war constituted the first systematic effort to apply and study the principles of proximity immediacy and expectancy (PIE) in the treatment of soldiers who developed CSR. These treatment principles, originally formulated during trench combat in World War I, are based on experience with soldiers who developed CSR and were removed from the artillery range. The basic premise of this treatment approach (which was never tested with appropriate scientific tools) is that the chances of overcoming CSR are highest when treatment is applied in proximity to the war zone, as early as possible after the condition develops, with an expectancy for full and rapid return to functioning. These relatively simple treatment principles were applied by the medical teams, including psychologists and social workers, that were part of the fighting units and were specially trained to deliver this treatment. Due to constraints of the war situation, the treatment was not available to all soldiers who developed CSR, and thus the prospective short-term and long-term outcome of treated and untreated veterans with CSR was limited. The data that emerged from these studies indicate that (1) CSR constituted approximately 15–25% of war-related casualties; (2) only one-half of the veterans who eventually developed PTSD applied for treatment in the front line; (3) among those who were treated according to the PIE principles, 15–20% eventually developed PTSD; and (4) the long-term beneficial effects of the PIE treatment need to be further explored.

Prisoners of War

Prisoners of war (POWs) from these wars were also studied. Twenty years after their captivity, 13% of the POWs still suffered from PTSD, compared to only 3% in a comparison group of combat veterans. It was also evident that recovery from PTSD among POWs was slower and incomplete compared to the control group.

Secondary Traumatization

Life often becomes literally a nightmare for many of those suffering from PTSD. What has been realized only in more recent years is the impact of the disorder on those living with the victims and on their caregivers. It has been shown that both close family members and therapists who are exposed to the traumatic memories and suffering of the victims tend to develop emotional distress and PTSD-like symptoms, including intrusive thoughts, emotional numbing and avoidance, and symptoms of increased arousal. This phenomenon was called secondary traumatization.

Traumatic Injury and Posttraumatic Stress Disorder

Another issue that received attention during the last years of the war activity in Lebanon, before the final withdrawal of the Israeli army from Lebanon, was the effect of traumatic injury on the development of PTSD. The assumption that bodily injury inflicted during the trauma does not increase the risk for PTSD, and in fact might even serve as a protective factor, prevailed among trauma experts in the past. It was believed that, being visible, the injury engenders sympathy and support from the environment to a greater extent than the psychic wounds of the trauma. It was further postulated that the injury also provides a focus for coping that distracts the subject's mind from the trauma and, last but not least, in the case of war, evacuation due to the injury removes the source of anxiety and danger from the soldier by providing a legitimate escape from the stressful situation. However, when injured and uninjured soldiers who were involved in identical war activities in Lebanon were compared for posttraumatic symptoms, it was found that the injured soldiers had an eightfold higher risk for developing PTSD compared to their uninjured peers. These findings clearly demonstrate that injury

in the context of trauma exerts an additional risk factor and increases the risk of PTSD. Possible explanations for this finding are that injury amplifies the sense of horror and perceived threat for life. In addition, the medical procedures and the pain resulting from the injury create an additional source of stress and serve as retrieval cues for the traumatic memories.

Lessons from the Persian Gulf War

During the Persian Gulf War in 1991, the entire population in Israel was, for the first time, under a direct threat of warfare and missile attacks. This constituted a serious psychological threat to the civilian population that had not been directly involved in war activity until that time. Data are not available about the prevalence and the magnitude of acute traumatic stress reactions in the entire population during that period. However, data are available about those who were evacuated to hospitals following the missile attacks.

Over a period of 2 weeks in the winter of 1991, approximately 40 surface-to-surface Iraqi SCUD missiles were launched and sent to Israel, targeting mainly the central coastal region of the greater Tel Aviv area. The uncertainty about the place, time, and type of warhead (chemical, biological, or conventional) was a source of stress and anxiety for the entire population, mainly in the targeted region. Furthermore, it was the cause of many traumatic stress reactions at or near the sites where the missiles hit. Approximately 800 citizens were evacuated to regional hospitals following these attacks, of which approximately 40% were diagnosed as psychological casualties. In addition, approximately 30% of the evacuees had mistakenly injected themselves with atropine, an antidote used during chemical warfare. These findings indicate that psychological responses constitute a significant portion of the casualties resulting from this kind of warfare and that rescue teams and hospitals have to be trained and prepared accordingly.

The specific psychological responses of mothers and children following the missile attacks were also studied. In a survey done 2 years after the war, families that were evacuated from homes that were destroyed or damaged in the missile attack were compared to families who did not have to leave their homes and were not directly affected by missile attacks. Mothers and children from the evacuated families manifested more posttraumatic symptoms and higher degrees of emotional distress. Stress symptoms tended to decrease over time in the displaced children from the evacuated families but not in their mothers. No differences in the children's adaptive behavior in the two groups were observed. Mothers who manifested more posttraumatic symptoms during the war were found to be more symptomatic at the follow-up, and the degree of stress in the children correlated with the mother's symptoms. These findings suggest that a maternal stress-buffering capacity is crucial for children's effective coping with stressful situations. The degree of family cohesion was also found to be an important factor in the successful coping of these children with the stressful situation. A follow-up study 5 years after the Gulf War showed significant decrease in posttraumatic symptoms in the evacuated children.

The Impact of Terror on Traumatic Stress Responses in the Israeli Population

The years 1987–2005 of the Israeli-Palestinians conflict, known as the Intifada, have exposed the civilian population in Israel to an unprecedented degree of violence and direct threat to lives as a result of numerous terror attacks and suicide bombers. The peak of the terror wave was in the years 2000–2004, and during this period 963 Israelis were killed and more than 6300 were wounded in hundreds of terror attacks and many more were directly exposed to the impact of terror and violence with direct threat to their lives.

A survey conducted in 2002 revealed that 16% of a representative population sample of approximately 560 people had been directly exposed to a terrorist attack and 37% had a family member or a friend who had been exposed. Ten percent of the sample met the diagnostic criteria for PTSD and approximately 60% reported feeling depressed. However, approximately 80% of the respondents expressed optimism about their personal future and 67% were optimistic about the future of the country. Seventy-five percent of the respondents expressed self-confidence about their ability to function in a terrorist attack. Despite the fact that many of the respondents expressed a low sense of safety with respect to themselves and their relatives, only 5% reported a need for professional help. Altogether, these findings indicate that, considering the nature and length of the Israeli traumatic experience during this period, the psychological impact of the situation may be considered moderate to low, and despite the fact that participants showed distress and a lowered sense of safety, only a minority developed psychiatric distress. These apparently discrepant findings may be related to a process of habituation and to effective psychological coping mechanisms.

Interesting differences were evident between urban communities exposed to different levels of stress and terror activities. To a certain degree, those communities

that were exposed to more violence and terror activity showed better coping abilities and lower degrees of personal distress than did communities with low exposure when facing stressful conditions.

Gender differences in the prevalence of PTSD have been reported. Although men are more frequently exposed to traumatic events, most studies have reported higher rates of PTSD among women. Studies done in Israel during 1999–2003 showed lower rates of direct exposure to terror among women, yet, compared to men, women reported more subjective distress and more posttraumatic symptoms and were using more coping strategies to overcome the distress. Sixteen percent of the women (but only 2.5% of the men) in the Israeli survey met the full criteria for PTSD.

Neurobiology and Psychophysiology of Posttraumatic Stress Disorder

Sleep and Posttraumatic Disorder

Sleep complaints are among the most frequently reported symptoms in PTSD. Difficulty falling asleep, restless and interrupted sleep, and recurrent dreams with nightmares are very common complaints.

Sleep disturbances and complaints of disrupted sleep in the early aftermath of the trauma have been associated with higher rates of PTSD later on. Some people with PTSD tend to develop a reversed sleep cycle, staying wake most night hours with a tendency to oversleep during the day. However, despite the high prevalence of subjective sleep complaints in PTSD, objective sleep measures have shown inconsistent findings. Unlike major depression, in which shortened rapid eye movement (REM) sleep latency is a well-established finding, polysomnographic sleep studies in patients with PTSD have not defined a clear pattern of sleep alterations that might be correlated with the subjective complaints of the patients. Some studies have shown fragmented REM sleep, whereas others have shown reduction in slow-wave sleep or increased awakening thresholds from slow-wave sleep. However, these findings are inconsistent. It has been suggested by some researchers that the subjective complaints of disrupted sleep in PTSD that are not corroborated in the polysomnographic studies are due to an altered sleep perception, namely, that the neurophysiology of sleep is not disturbed but there is a state of sleep misperception that causes a dissociation between an apparently normal sleep pattern and its altered perception by the subject. This could explain the relative inefficacy of sleep medications in PTSD, which is a widely acknowledged phenomenon and results in the excessive use and often abuse of sleep medications in these cases.

Physiological Arousal and Posttraumatic Disorder

Symptoms of increased arousal and exaggerated startle response are characteristic of PTSD. When exposed to loud tones 1 and 4 months after a traumatic event, subjects who developed PTSD showed a greater heart rate response, altered skin conductance, and eye muscle electromyogram (EMG) responses when compared to subjects who did not develop PTSD after a traumatic event. No such differences were evident between the groups immediately after the trauma, indicating that these findings were state-dependent and characteristic of trauma survivors who develop PTSD.

Furthermore, subjects with PTSD failed to show the habituation in response to repeated loud tones that is normally seen and was evident in the comparison subjects who did not develop PTSD. This phenomenon could reflect a more generalized inability for habituation of stressful signals; the recurrent intrusive memories in PTSD could be considered to be a failure of habituation to the content of the traumatic memories.

Heart Rate and Heart Rate Variability in Posttraumatic Disorder

Increased heart rate is a determinant of physiological arousal seen in PTSD. Subjects who develop PTSD have been shown to have higher heart rates during the immediate aftermath of the trauma compared to traumatized individuals who do not develop PTSD. The elevated heart rate is caused by the increased noradrenergic tone, suggesting that increased noradrenergic activity immediately after the trauma may play an important role in the neurobiological processes involved in the development of PTSD. From a clinical perspective, this finding suggests that elevated heart rate immediately after the trauma is a predictor of PTSD.

Animal Model for Posttraumatic Stress Disorder

The validity of an animal model for psychiatric disorders is often questionable. However, there is a clear advantage for PTSD because an exposure to threat (a predator, in the case of a rat) certainly addresses some of the validity issues that might be problematic in animal models of other psychiatric disorders. An animal model of PTSD composed of exposure to a predator (the smell of a cat, in this case), applying behavioral cut-off criteria to differentiate between the maladaptive rats and the well-adapted rats was developed in Israel in 2003. This enables investigators to further advance PTSD research by examining

questions such as early life adversities, stress-dependent increase in cortisol levels, and therapeutic (preventive) effects of early administration of medications that are often used in the treatment of PTSD, such as selective serotonin reuptake inhibitors (SSRIs).

Conclusion

The unique (although unfortunate) combination in Israel of having many Holocaust survivors, several wars, exposure to repeated waves of terror, and the presence of a pure form of PTSD (i.e., very low rates of alcoholism and drug abuse), coupled with the expertise in research, have helped Israeli researchers and clinicians to develop vital and advanced treatment and research in stress-related disorders, including acute stress reaction, CSR, and PTSD. The key findings have called into question some of the accepted notions of PTSD. Issues such as a direct linkage between terror exposure and PTSD were not found. The use of PIE principles in the modern battlefield seems to require further research, as do the apparently strong link between physical injury and PTSD and the lack of objective findings in a sleep laboratory in regard to sleep disturbance in PTSD. The development of a new animal model for PTSD, with the focus on behavioral cut-off criteria, along with further research in regard to the genetic vulnerability to PTSD, may pave ways to deepen our understanding of this intriguing phenomenon.

Further Reading

Barak, Y., Aizenberg, D., Szor, H., et al. (2005). Increased risk of attempted suicide among aging holocaust survivors. *American Journal of Geriatric Psychiatry* **13**(8), 701–704.

Bleich, A., Gelkopf, M. and Solomon, Z. (2003). Exposure to terrorism, stress-related mental health symptoms, and coping behaviors among a nationally representative sample in Israel. *Journal of the American Medical Association* **290**(5), 612–620.

Cohen, H. and Zohar, J. (2004). An animal model of posttraumatic stress disorder: the use of cut-off behavioral criteria (review). *Annals of the New York Academy of Sciences* **1032**, 167–178.

Gil, S., Caspi, Y., Ben-Ari, I. Z., et al. (2005). Does memory of a traumatic event increase the risk for posttraumatic stress disorder in patients with traumatic brain injury?: a prospective study. *American Journal of Psychiatry* **162** (5), 963–969.

Koren, D., Arnon, I. and Klein, E. (1999). Acute stress response and posttraumatic stress disorder in traffic accident victims: a one-year prospective, follow-up study. *American Journal of Psychiatry* **156**(3), 367–373.

Koren, D., Arnon, I., Lavie, P., et al. (2002). Sleep complaints as early predictors of posttraumatic stress disorder: a 1-year prospective study of injured survivors of motor vehicle accidents. *American Journal of Psychiatry* **159**(5), 855–857.

Koren, D., Norman, D., Cohen, A., et al. (2005). Increased PTSD risk with combat-related injury: a matched comparison study of injured and uninjured soldiers experiencing the same combat events. *American Journal of Psychiatry* **162**(2), 276–282.

Laor, N., Wolmer, L. and Cohen, D. J. (2001). Mothers' functioning and children's symptoms 5 years after a SCUD missile attack. *American Journal of Psychiatry* **158**(7), 1020–1026.

Laor, N., Wolmer, L., Mayes, L. C., et al. (1997). Israeli preschool children under Scuds: a 30-month follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry* **36**(3), 349–356.

Nadler, A. and Ben-Shushan, D. (1989). Forty years later: long-term consequences of massive traumatization as manifested by Holocaust survivors from the city and the kibbutz. *Journal of Consulting and Clinical Psychology* **57**(2), 287–293.

Pillar, G., Malhotra, A. and Lavie, P. (2000). Post-traumatic stress disorder and sleep – what a nightmare! *Sleep Medicine Reviews* **4**(2), 183–200.

Shalev, A. Y. and Freedman, S. (2005). PTSD following terrorist attacks: a prospective evaluation. *American Journal of Psychiatry* **162**(6), 1188–1191.

Shalev, A. Y., Freedman, S., Peri, T., et al. (1998). Prospective study of posttraumatic stress disorder and depression following trauma. *American Journal of Psychiatry* **155** (5), 630–637.

Shalev, A. Y., Peri, T., Brandes, D., et al. (2000). Auditory startle response in trauma survivors with posttraumatic stress disorder: a prospective study. *American Journal of Psychiatry* **157**(2), 255–261.

Solomon, Z., Shklar, R. and Mikulincer, M. (2005). Frontline treatment of combat stress reaction: a 20-year longitudinal evaluation study. *American Journal of Psychiatry* **162**(12), 2309–2314.

Solomon, Z. and Lavi, T. (2005). Israeli youth in the Second Intifada: PTSD and future orientation. *Journal of the American Academy of Child and Adolescent Psychiatry* **44**(11), 1167–1175.

Solomon, Z., Neria, Y., Ohry, A., et al. (1994). PTSD among Israeli former prisoners of war and soldiers with combat stress reaction: a longitudinal study. *American Journal of Psychiatry* **151**(4), 554–559.

Trier Social Stress Test

B M Kudielka and S Wüst
University of Trier, Trier, Germany
C Kirschbaum
Technical University of Dresden, Dresden, Germany
D H Hellhammer
University of Trier, Trier, Germany

© 2007 Elsevier Inc. All rights reserved.

Introduction: Origin of the Trier Social Stress Test (TSST)
Description of the TSST Protocol
Overview of Sources of Intra- and Interindividual Differences
The TSST in Clinical Populations
Conclusions

Glossary

Cortisol — A glucocorticoid hormone produced by the adrenal cortex. Cortisol exhibits a pronounced circadian rhythm and responds to a wide range of stressors. It has many physiological effects and virtually every nucleated cell in the body has cortisol receptors. It is predominantly (90–95%) bound to proteins in blood; only 5–10% of the total plasma cortisol circulates biologically active as unbound, free cortisol.

Hypothalamic-pituitary-adrenocortical (HPA) axis — A central regulatory and control system of the organism that connects the central nervous system with the endocrine system. Under stress, the hypothalamus secretes corticotropin-releasing hormone, which provokes the release of adrenocorticotropic hormone (ACTH) from the pituitary. ACTH triggers the secretion of cortisol from the adrenal cortex. The functioning of the axis is controlled by several negative feedback loops.

Psychosocial stress — Psychological stress that exerts physiological, behavioral, and/or verbal-subjective stress responses as a result of stressful social interactions, which often involve social-evaluative threat (in contrast to stress responses provoked by physical exertion or pharmacological stimulation).

Social-evaluative threat — Social component of a psychological or a psychosocial stress task in which task performance could be negatively judged by others. Stress tasks containing the two components uncontrollability and social-evaluative threat are associated with the largest hypothalamic-pituitary-adrenocortical axis stress responses and the longest recovery times.

Sympathetic-adrenal-medullary (SAM) axis — Sympathetic nerves, originating in the central nervous system, stimulate the adrenal medulla, which in turn secretes epinephrine and norepinephrine. These catecholamines mobilize energy-producing mechanisms quickly and downregulate less important organ functions.

Trier social stress test (TSST) — A standardized stress protocol for the induction of moderate psychosocial stress in laboratory settings, which mainly consists of a brief preparation period (3 min) and a test period in which the subject has to deliver a free speech (5 min) and perform mental arithmetic (5 min) in front of an audience. The TSST can be applied to younger and older adults, to children as well as clinical populations. The TSST is one of the few available stress protocols that satisfies the criteria of a motivated performance task that combines elements of uncontrollability and high levels of social-evaluative threat.

Introduction: Origin of the Trier Social Stress Test (TSST)

Stress is one of the most significant health problems in the twenty-first century. This explains a pressing need for investigations into the biological pathways linking stress and health. Besides the sympathetic-adrenal-medullary (SAM) axis, the hypothalamic-pituitary-adrenocortical (HPA) axis is the major physiological stress response systems in the organism. Since alterations in HPA axis stress responses appear to be a close correlate or even a determining factor of the onset of different diseases or disease progression, the characterization of an individual's HPA axis response pattern to psychosocial stress appears to be of major interest. For such research agenda, a novel research tool was needed since available laboratory stress protocols yielded insufficient activation of the HPA axis. A stress protocol that was potent enough to induce significant changes of endocrine and cardiovascular parameters in the majority of subjects tested was eventually developed and termed the Trier Social Stress Test (TSST). Over the past decade, well over 4000 TSST sessions were performed in many laboratories worldwide. Thus, the TSST has become a standard protocol for the experimental induction of psychological stress

in healthy subjects as well as in clinical populations, investigating a wide range of different outcome variables ranging from subjective-verbal stress reports to objective behavioral and biological stress responses, including parameters of the HPA and SAM axis and the cardiovascular, immunological, and blood coagulation system.

Recently, Dickerson and Kemeny conducted a meta-analysis based on 208 laboratory studies of acute psychological stressors in order to delineate the essential situational elements capable of eliciting HPA axis responses. The results show that motivated performance tasks elicit cortisol responses if they are (1) uncontrollable, (2) creating a context of forced failure in which a subject is unable to avoid negative consequences or cannot succeed despite his or her best effort, or (3) characterized by a social component called social-evaluative threat, in which task performance could be negatively judged by others. Tasks containing both components (uncontrollability and social-evaluative threat) were associated with the largest HPA axis stress responses and the longest recovery times. These meta-analytical findings fit to the theoretical reasoning (based on the social self-preservation theory) that "uncontrollable threats to the goal of maintaining the 'social self' would trigger reliable and substantial cortisol changes" (Dickerson and Kemeny, 2004). The authors concluded that the TSST is one of the few available stress protocols that satisfies the criteria of a motivated performance task that combines elements of uncontrollability and high levels of social-evaluative threat. Their results fit with earlier work by Mason, stating that whenever a situation (or a stimulus) is perceived as novel, uncontrollable, unpredictable, or ambivalent, or if the individual anticipates negative organismic or psychological consequences, the negative feedback signals from the hippocampus and other sites will be overridden, resulting in a cascade of corticotropin-releasing hormone (CRH), adrenocorticotropic hormone (ACTH), and cortisol secretion.

Description of the TSST Protocol

The TSST is a motivated performance task consisting of a brief preparation period (3 min) followed by a test period in which the subject has to deliver a free speech (5 min) and perform mental arithmetic (5 min) in front of an audience. The total exposure time adds up to 13 min. The following description of the TSST procedure primarily applies to investigations of HPA axis responses, and, therefore, other outcome measures may require some modifications.

The test onset should be preceded by a rest period of at least 45 min to minimize unrelated short-term effects on the outcome measures. The subject is asked to take over the role of a job applicant who is invited for a personal interview with the company's staff managers (selection committee). The selection panel members are introduced as being specially trained to monitor nonverbal behavior, and subjects are informed to expect tape and video recordings for later analysis of their performance. The selection committee is composed of two or three male and female collaborates of the experimenter, unacquainted with the subject. The committee is trained to communicate with the subject in an unresponsive neutral manner and does not respond with any facial or verbal feedback (note: no harassment or anger induction). The committee first requests that the subject sit down (not if blood pressure is measured) at a small table equipped with paper and pencil to prepare the talk (preparation period). After 3 min, the subject is asked to step forward to a marked spot in front of the microphone about 1–3 m apart from the committee's table and to start the oral presentation (note: the use of written concept is not allowed). Whenever a subject finishes the speech in less than 5 min, the managers respond in a standardized way. First they tell the volunteer, "You still have some time left. Please continue." Should the subject finish a second time before the 5 min are over, the managers are quiet for 20 s and then ask prepared questions (e.g., What are your personal strengths? What are your major shortcomings? Do you have enemies? Why? What do you think about team work?). After 5 min, the selection committee interrupts and asks the subject to serially subtract the number 17 from 2023 as fast and as accurately as possible. On every failure the subject has to restart at 2023 with one member of the committee interfering "Stop – mistake – start over at 2023, please." After 5 min the task is terminated.

Samples are to be gained before onset and after cessation of the stress task depending on the dynamic of the selected outcome variables and should cover prestress levels, the initial stress response, peak level, and recovery. Prestress HPA axis hormone levels should not be interpreted as baseline levels since they are substantially affected by the subjects' anticipation of the upcoming challenge. At the end of the test session the subject will be fully debriefed about the goal of the study and the nature of the stressor. A full debriefing includes a visit of the collaborates who took over the role of the selection committee.

With this protocol, salivary cortisol levels rise two- to threefold in about 70–80% of all subjects, with peak levels around 10–20 min after cessation of the stress task. In addition to free cortisol, the levels of total plasma cortisol, ACTH, catecholamines (epinephrine, norepinephrine), growth hormone, prolactin,

testosterone, systolic and diastolic blood pressure, heart rate, immune parameters (e.g., neutrophils, eosinophils, basophils, lymphocytes, interleukin-6, TNFα), glucose, α-amylase, endogenous inhibitor of monoamine oxidase (MAO-AI), as well as measures of hemoconcentration (e.g., hematocrit, haemoglobin, plasma volume) and blood coagulation (fibrinogen, von Willebrand factor antigen, D-dimer, clotting factors) increase significantly.

The TSST protocol can also be employed to test children and older adults after changing the instructions slightly for the respective age group. In elderly adults, subjects are instructed to apply for a (voluntary) part-time job (e.g., child caring, housekeeping, technical assistant). To help a retired subject take over the role of a job applicant, a fabricated newspaper advertisement can be presented (tested age range 59–91 years). The adapted TSST for children (TSST-C) also consists of a preparation period, public speaking, and mental arithmetic task. In the speaking part, children receive the beginning of a story and are told that they should finish telling the story, making it as exciting as possible, in front of a jury (tested age range: 7–14 years). The numbers of the mental arithmetic task are adapted to the performance level of the respective age group.

Although areas under the response curve (AUCs) are much higher in the morning due to the circadian rhythm of HPA axis activity, comparable free cortisol net increases after TSST exposure can be assessed with equal reliability in the morning and afternoon. However, in early morning sessions the experimenter should ensure that the onset of a stress experiment does not interfere with the morning cortisol awakening rise (CAR). Also, the appropriate time window for test sessions in the morning is smaller than in the late afternoon.

Overview of Sources of Intra- and Interindividual Differences

This section summarizes findings from studies investigating various factors that potentially contribute to differences in HPA axis responses to the TSST. One of the most prominent features of HPA axis (and other system) responses after psychological stress is the large variation in response magnitude among subjects and across test sessions. Significant differences can be observed with respect to the net hormone output as well as the time course of hormone secretion. Considerable evidence has accumulated for a significant impact of a variety of moderating and intervening variables, which are briefly summarized in the following.

Gender and Sex Steroids

One of the most consistent findings from employing the TSST is the significantly larger salivary cortisol and ACTH response (up to twice as high) in men compared to women, while prestress levels are not considerably different. The same sex effect emerges for elderly subjects (see also below). The observed sex differences point to a possible impact of female sex steroids as potential mediators of HPA axis responses to psychosocial stress. In a study of 81 healthy young adults, the TSST was applied to men, women in the follicular phase of the menstrual cycle, women in the luteal phase, and women using oral contraceptives. The results disclosed that significant sex differences emerge for ACTH and free salivary cortisol but not for total plasma cortisol stress responses. Moreover, ACTH responses were elevated in men compared to women, regardless of menstrual cycle phase or use of oral contraceptives. Women in the luteal phase had saliva cortisol stress responses comparable to men, whereas women in the follicular phase or women taking oral contraceptives showed significantly lower free cortisol responses. These observations point to the necessity of strictly distinguishing between the total cortisol secretion and the levels of bioavailable free cortisol. Whereas many animal studies can be cited that have directly investigated the impact of estrogens on HPA axis regulation, few experimental studies have been conducted in humans, and the empirical evidence is rather inconsistent.

Age

A recent reanalysis of five independent TSST studies based on 102 healthy subjects between 9 and 76 years showed that the stress task induced significant HPA axis responses in all age groups in both males and females. The data revealed no sex differences in free cortisol responses in children and younger adults, but larger free cortisol responses in elderly men than in elderly women. This effect did not appear to be attributable to self-reported stress level differences. For total plasma cortisol, elderly women showed generally enhanced cortisol response levels compared to elderly and younger men as well as younger women. For ACTH, the response was higher in younger adults, primarily due to an elevated response in younger men.

With respect to autonomic responses, differential heart rate responses and recovery after exposure to the TSST were observed in healthy children, younger adults, and elderly adults. Aggregated data across the five studies showed that there was a marked age effect in the heart rate stress response, with children and younger adults showing significantly higher increases

than elderly persons. This effect is probably caused by well-known age-related alterations in heart rate regulation. The analysis of sex effects showed that girls had higher heart rate increases during the stress exposure than boys. In younger adults, stress responses were also higher in women. Peak heart rate responses were comparable in older men and women, with only men returning to prestressor baselines during the observation period.

Lactation and Breastfeeding

Lactation in women, in contrast to rats, does not result in a general restraint of HPA axis responses to a psychosocial stressor such as the TSST. Rather, suckling seems to exert a short-term suppression of the cortisol response to psychosocial stress.

Nicotine, Coffee, Alcohol Consumption, and Dietary Energy Supplies

Nicotine is a potent stimulator of the HPA axis through induction of CRH release after binding to cholinergic receptors in the locus coeruleus and hypothalamus, and chronic smoking changes the HPA axis responses to stress. Blunted cortisol responses to the TSST were repeatedly observed in habitual smokers. It has also been shown that caffeine intake can activate important components of the pituitary-adrenocortical response in humans during resting states, and there is some evidence for combined stimulatory effects of caffeine and psychological stress on HPA axis responses. Relationships between alcohol and acute stress can be at least threefold: (1) alcohol intake can alter physiological responses to acute stress, (2) acute social stress can impact on subsequent net alcohol consumption in healthy social drinkers, and (3) acute HPA axis stress responses as well as effects of acute alcohol intake before the TSST can be altered in subjects with a positive family history of alcoholism. Finally, the availability of dietary energy supplies (e.g., short-term fasting, glucose availability) appears to exert important regulatory functions in pituitary-adrenal stress responses, pointing to an important role of the nutritional state.

Social Support and Social Hierarchy

Few data provide experimental support for the hypothesis that social support might attenuate the HPA axis response to acute stress; however, effects appear to be sex specific. The neuropeptide oxytocin seems to enhance the buffering effect of social support on stress responsiveness, at least in men. These results concur with data from animal research suggesting an important role of oxytocin as an underlying biological mechanism for stress-protective effects of positive social support. Also, the position in a social hierarchy seems to be relevant for acute HPA axis stress responses. After the TSST, salivary cortisol levels highly increased in socially dominant subjects, while only a modest elevation was observed in subordinate men.

Personality Factors

It is tempting to speculate that the HPA axis response to stress is influenced, at least in part, by stable trait factors since the endocrine response to psychosocial stress can be viewed as a close interaction between situation and person variables within a given context. Interestingly, relationships between personality traits and stress responses only emerge after repeated exposures to the TSST. While novelty may mask the impact of personality on the cortisol stress response on the first exposure, differences in the ability to cope with the stressful situation may lead to different cortisol stress response patterns on subsequent stress exposures. With decreasing novelty and unpredictability, the impact of psychological trait variables increases, and a closer correlation with the endocrine parameters can emerge. Thus, a reliable investigation of associations between personality variables and cortisol stress responses appears to require repeated TSST exposure and data aggregation.

Habituation

After repeated exposure to (initially) stressful situations, a rapid habituation of HPA axis stress responses is normally observed. This has been reported consistently for the TSST. Recently, substantial interindividual variability of such habituation patterns was documented in a large study sample. While 52% of subjects showed the well-known HPA axis response habituation, almost 16% of the participants showed a response sensitization across three test sessions. Interestingly, altered HPA axis response habituation to TSST exposures seems to be associated with vital exhaustion, pointing to a potential mechanism for how exhaustion relates to increased disease vulnerability. However, it can be speculated that habituation might be rather specific for HPA axis parameters since TSST studies investigating stress-related changes in parameters of the sympathetic nervous system, blood coagulation indices, hemoconcentration, blood cells, or interleukin-6 (IL-6) levels did not report on habituation effects.

Interventions

First evidence shows that brief group-based cognitive-behavioral stress management training can reduce neuroendocrine or cardiovascular TSST stress

responses. It may thus be speculated that such interventions may prove useful in preventing detrimental stress effects in healthy subjects as well as patients. Other relaxation techniques (muscle relaxation, relaxing music) might also alter stress responses or facilitate recovery from TSST stress exposure.

Genetic Factors: Heritability and Polymorphisms

In addition to behavioral and situational variation, HPA axis activity appears to be profoundly influenced by genetic factors, as shown in quantitative genetic studies (family and twin studies) as well as candidate gene studies on polymorphisms in the genes coding for the glucocorticoid receptor, the mineralocorticoid receptor, the melanocortin-2 (ACTH) receptor, the melanocortin-4 receptor, the μ-opioid receptor, the $GABA_A$ (gamma-aminobutyric-$acid_A$) receptor, the α-adrenergic receptor, the $5\text{-}HT_{2A}$ ($serotonin_{2A}$) receptor, COMT (catechol-O-methyltransferase), TNFα (tumor necrosis factor α), and ACE (angiotensin-I converting enzyme).

TSST and Cellular Mechanisms

In a series of experiments applying the TSST stress protocol, Rohleder and co-workers investigated the influence of age, sex, and sex hormone status on glucocorticoid (GC) sensitivity of peripheral leukocytes since a cell's response to cortisol is predominantly determined by both the steroid level it is exposed to and its sensitivity for glucocorticoids, that is, the efficacy of glucocorticoid-mediated signal transduction.

Bierhaus et al. employed the TSST to elucidate how neuroendocrine stress responses can mediate cellular processes, which, in the long run, can promote vascular diseases. Their data provide convincing evidence that the activation of the transcription factor NF-κB (nuclear factor-κB), a protein that is sensitive to psychosocial stress, might represent a norepinephrine-dependent pathway converting psychosocial stress into cellular dysfunction.

The TSST in Clinical Populations

In addition to healthy study participants, the TSST can be used in clinical populations ranging from psychiatric diseases to somatic complaints as shown in several studies over the last few years. To date, the TSST has been successfully applied in patients suffering from, e.g., major depression, anxiety disorder, social phobia, posttraumatic stress disorder (PTSD), burnout, exhaustion, chronic fatigue syndrome, fibromyalgia, recurrent abdominal pain, chronic pelvic pain, temporomandibular dysfunction (myofascial pain and dysfunction), functional gastrointestinal disorders (irritable bowel syndrome or non-ulcer dyspepsia), different manifestations of chronic atopy, and atopic dermatitis or chronic allergic asthma, and adults with a history of sexual or physical childhood abuse or breast cancer survivors.

Conclusions

Over more than a decade of research, the TSST has proven a useful tool in the field of basic, applied, as well as clinical psychobiological research with a wide range of psychobiological outcome variables. As well as being used in the direct study of different biological stress systems like the HPA and SAM axis or cardiovascular system, the TSST can be used to investigate biological pathways between stress and disease or to study stress-related phenomena (e.g., effects of stress on memory and cognition) and interactions between drug and stress effects. Clinical studies continue to accumulate evidence that different diseases are associated with characteristic stress response profiles. Furthermore, biological mechanisms begin to unfold that could be helpful in explaining the stress–disease association. In a next step, the challenge for researchers will be to investigate whether the TSST protocol can be applied as a diagnostic tool for the prediction of disease susceptibility and symptom severity and/or for monitoring the efficacy of interventions.

See Also the Following Articles

Acute Stress Response: Experimental; Adrenal Cortex; Adrenal Medulla; Adrenaline; Adrenocorticotropic Hormone (ACTH); Aging and Psychological Stress; Aging and Stress, Biology of; Alarm Phase and General Adaptation Syndrome; Allostasis and Allostatic Load; Amygdala; Animal Models (Nonprimate) for Human Stress; Autonomic Nervous System; Blood Pressure; Cardiovascular System and Stress; Catecholamines; Central Stress Neurocircuits; Control and Stress; Coping Skills; Coping and Stress: A Lens and Filter Model; Corticosteroids and Stress; Corticotropin Releasing Factor (CRF); Endocrine Systems; Gender and Stress; Genetic Factors and Stress; Glucocorticoid Negative Feedback; Glucocorticoids, Effects of Stress on; Glucocorticoids, Overview; Heart Rate; Hypothalamic-Pituitary-Adrenal; Immune Cell Distribution, Effects of Stress on; Immune Function, Stress-Induced Enhancement; Immune Response; Learning and Memory, Effects of Stress on; Memory and Stress; Menopause and Stress; Neuroendocrine Systems; Neuroimmunomodulation; Psychological Stressors, Overview; Salivary Cortisol; Sex Differences in Human Stress Response; Social Stress, Animal Models of; Social Support; Stress Effects, Overview; Sympathetic Nervous System.

Further Reading

Bierhaus, A., Wolf, J., Andrassy, M., et al. (2003). A mechanism converting psychosocial stress into mononuclear cell activation. *Proceedings of the National Academy of Sciences USA* **100**(4), 1920–1925.

Biondi, M. and Picardi, A. (1999). Psychological stress and neuroendocrine function in humans: the last two decades of research. *Psychotherapy and Psychosomatics* **68**(3), 114–150.

Buske-Kirschbaum, A. and Hellhammer, D. H. (2003). Endocrine and immune responses to stress in chronic inflammatory skin disorders. *Annals of the New York Academy of Sciences* **992**, 231–240.

Buske-Kirschbaum, A., Jobst, S., Wustmans, A., et al. (1997). Attenuated free cortisol response to psychosocial stress in children with atopic dermatitis. *Psychosomatic Medicine* **59**(4), 419–426.

Dickerson, S. S. and Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychology Bulletin* **130**(3), 355–391.

Heim, C., Ehlert, U. and Hellhammer, D. H. (2000). The potential role of hypocortisolism in the pathophysiology of stress-related bodily disorders. *Psychoneuroendocrinology* **25**(1), 1–35.

Heinrichs, M., Neumann, I. and Ehlert, U. (2002). Lactation and stress: protective effects of breast-feeding in humans. *Stress* **5**(3), 195–203.

Kirschbaum, C., Pirke, K. M. and Hellhammer, D. H. (1993). The 'Trier Social Stress Test'—a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* **28**(1–2), 76–81.

Kirschbaum, C., Kudielka, B. M., Gaab, J., Schommer, N. C. and Hellhammer, D. H. (1999). Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosomatic Medicine* **61**(2), 154–162.

Kudielka, B. M. and Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. *Biological Psychology* **69**(1), 113–132.

Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H. and Kirschbaum, C. (2004). Differential heart rate reactivity and recovery after psychosocial stress (TSST) in healthy children, younger adults, and elderly adults: the impact of age and gender. *International Journal of Behavioral Medicine* **11**(2), 116–121.

Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H. and Kirschbaum, C. (2004). HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: impact of age and gender. *Psychoneuroendocrinology* **29**(1), 83–98.

Kudielka, B. M., Schommer, N. C., Hellhammer, D. H. and Kirschbaum, C. (2004). Acute HPA axis responses, heart rate, and mood changes to psychosocial stress (TSST) in humans at different times of day. *Psychoneuroendocrinology* **29**(8), 983–992.

Kudielka, B. M., Hellhammer, D. H. and Kirschbaum, C. (In press). Ten years of research with the Trier Social Stress Test (TSST) – revisited. In: Harmon-Jones, E. & Winkielman, P. (eds.) *Social Neuroscience*. New York: Guilford Press.

Mason, J. W. (1968). A review of psychoendocrine research on the pituitary-adrenal cortical system. *Psychosomatic Medicine* **30**(Supplement 5), 576–607.

Rohleder, N., Wolf, J. M. and Kirschbaum, C. (2003). Glucocorticoid sensitivity in humans—interindividual differences and acute stress effects. *Stress* **6**(3), 207–222.

Wüst, S., Rossum, E. F. C., Koper, J. W., Kumsta, R., Entringer, S. and Hellhammer, D. H. (2004). A psychobiological perspective on genetic determinants of hypothalamus-pituitary-adrenal axis activity. *Annals of the New York Academy of Sciences* **1032**, 52–62.

Wüst, S., Federenko, I. S., van Rossum, E. F., Koper, J. W. and Hellhammer, D. H. (2005). Habituation of cortisol responses to repeated psychosocial stress—further characterization and impact of genetic factors. *Psychoneuroendocrinology* **30**(2), 199–211.

Tuberculosis *See:* Macrophage Antimycobacterial Activity, Effects of Stress on.

Type A Personality, Type B Personality

W S Shaw
University of Massachusetts, Worcester, MA, USA
J E Dimsdale
University of California, San Diego, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by W S Shaw and J E Dimsdale, volume 3, pp 624–630, © 2000, Elsevier Inc.

Introduction
Historical Perspective
Conceptualization and Assessment
Empirical Support
Component Factors
Beyond Coronary Heart Disease
Conclusion

Glossary

Free-floating hostility	Behavior characterized by hostile vocal inflection and facial expression, ill temper, cynicism, sleeplessness, irritability, and interpersonal conflict.
Structured interview	An assessment technique involving a series of questions meant to elicit impatience, hostility, and competitiveness among those believed to have type A-defining characteristics.
Time urgency	Behavior characterized by extreme haste, punctuality, time monitoring, and tendencies to engage in polyphasic activities with poor recall of events.
Type A behavior	"An action-emotion complex that can be observed in any person who is aggressively involved in a chronic, incessant struggle to achieve more and more in less and less time, and if required to do so against the opposing efforts of other things and other persons" (Friedman and Rosenman, 1974: 67).
Type B behavior	A relative absence of the behavioral characteristics defined as type A.

Introduction

For centuries, observers have suggested that certain personality patterns are toxic to the heart. Early reports were based on clinical observations and, as is typical in the history of medicine, such reports were either case reports or pronouncements (in lieu of data) by respected authorities. From 1930 to 1950, such reports were typically framed from a psychoanalytic perspective. However, shortly afterward, a new conceptualization emerged. The investigators claimed to have recognized a coronary-prone personality pattern but chose to label it in a value-neutral fashion as type A behavior pattern.

Historical Perspective

Few health-related psychological constructs have received the scientific interest and widespread public recognition of the type A personality classification. Scientific interest in type A and its relation to coronary heart disease began in the 1960s with a research team at the Mount Zion Hospital in San Francisco, California. Meyer Friedman, Ray Rosenman, and colleagues identified a constellation of personality traits that were observed disproportionately among their patients presenting with coronary heart disease (CHD). These traits (labeled type A or coronary-prone) included excessive drive, time urgency, ambition, impatience, aggressiveness, hostility, and competitiveness. When confronted with stressful situations, type A individuals were believed to react with rapid, explosive retorts and increased muscle tension. These behaviors were thought to cause heightened or excessive wear on the cardiovascular system and thereby increase the risk of CHD. The relative absence of these traits defined type B or non-coronary-prone behavior.

Clinical reports suggesting a coronary-prone behavior pattern led to a multitude of epidemiological and experimental studies in the 1970s and 1980s to examine these relationships in the general population and to elucidate biological mechanisms. Assessment techniques were developed, including structured interviews, self-report questionnaires, and behavioral observations. Studies examined the relationship of type A to CHD incidence, angina, myocardial infarction (MI), atherosclerosis, and cardiovascular reactivity. Early epidemiologic studies supported an association between type A and CHD, and this led to a 1981 consensus statement by the Review Panel on Coronary-Prone Behavior and Coronary Heart Disease (organized by the National Heart, Lung, and Blood Institute) listing type A as a viable risk factor for heart disease.

Despite the early support for type A, a proliferation of research in the 1980s failed to replicate earlier findings, and several related psychological constructs (hostility, anger, and aggression) emerged as stronger predictors of CHD. As a result, interest in type A has waned in recent years. By the late 1980s, two

meta-analyses concluded that the accumulated empirical studies were inconsistent and that prospective evidence linking type A to CHD was weak. Since that time, behavioral CHD research has emphasized hostility and other related constructs over the global type A characterization. However, originators of the type A concept continue to defend its role as a CHD risk factor, blaming poor assessment techniques and a substantial drift from its original conceptualization for the discrepant research findings.

As late as the 1990s, interest in the health implications of type A behavior continued, albeit in a more limited scope. Meyer Friedman and his colleagues at the University of California, San Francisco, continued to refine clinical assessment techniques and develop interventions for reducing type A behavior. Others sought to examine type A behavior with respect to other psychosomatic health problems. Today, there is declining research interest in type A as a behavioral health risk. Figure 1 illustrates the rapid rise and fall of this construct as a title word in psychological research publications over the past 40 years. Nevertheless, type A has served as an important catalyst for the explosion of scientific and public interest in behavioral factors associated with health, especially cardiovascular disease.

Conceptualization and Assessment

Considerable conceptual confusion has surrounded the type A construct. Type A behavior has been variously described as a medical disorder, an inappropriate coping mechanism, a strong need for productivity, or a pattern of physiological reactivity. Although most research has presumed type A behavior constitutes an enduring personality trait, other studies have attributed type A behavior to specific cognitions that can be modified through therapeutic interventions. Much of the accumulated discrepancy in results between studies of type A may be attributable to poor agreement in construct definition and assessment.

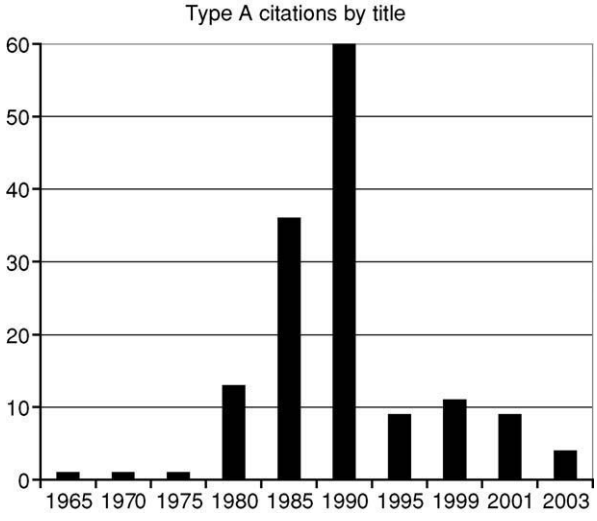


Figure 1 Type A journal title citations.

Type A behavior was first assessed through clinical ratings applied to structured interviews or videotaped structured interviews, and these interview techniques have been further updated and expanded by Meyer Friedman and colleagues. The current technique combines both observations and subjective reporting of patients using eliciting remarks. The interview consists of multiple queries that are designed to elicit manifestation of time urgency and free-floating hostility that emerge for type A, but not for type B, individuals. Questions include "Do you mind very much waiting in grocery checkout, bank, or theater lines or waiting to be seated in a restaurant?" (time urgency) and "Do you often find it difficult to fall asleep or to continue to sleep because you are upset about something a person has done?" (hostility). Scoring is based on both endorsements of items and the presence of psychomotor signs. For time urgency, signs include aspects of facial tension, posture, speech, breathing, and perspiration. For free-floating hostility, these include facial expression, eyelid movement, vocal orientation, and hand clenching. For each item, scores are clinical ratings varying in assigned weights and ranges. Total scores range from 0 to 480, and a total score greater than 45 indicates the presence of type A traits.

As an alternative to the interview technique, self-report measures of type A were developed to simplify data collection for use in large-scale epidemiological studies of CHD. One of the earliest and most frequently used questionnaires is the Jenkins Activity Scale (JAS). Other pen-and-paper measures of type A behavior include the Framingham Type A Scale, the Bortner Scale, the MMPI-2 Type A Scale, the Gough Adjective Checklist, the Thurstone Activity Checklist, the Coronary Prone Attitudes Scale, and the Ketterer Stress Symptoms Frequency Checklist.

The content and emphasis of type A self-report measures vary substantially. The JAS includes subscales of hard-driving competitiveness, speed and impatience, and job involvement. The Framingham Type A Scale includes time urgency, competitive drive, and perceived job pressures, but does not include hostility. The Bortner scale contains 14 adjective pairs between which respondents mark the level of agreement with either type A or type B traits (e.g., "Never late" versus "Casual about appointments"). The MMPI-2 Type A Scale includes 19 items that assess the type A components of hostility, competitiveness, and time urgency; these were selected from the

total pool of MMPI-2 questions by expert consensus and verified using empirical tests of item-to-type A group classification. The Ketterer Scale compares self-ratings of type A characteristics with parallel ratings of a friend or family member.

The reliability of the various interview and self-report measures is quite good. For example, interrater correspondence for discriminating type A and B individuals using the videotaped structured interview has ranged from 75 to 90%. For the MMPI-2 Type A Scale, test–retest reliability is high ($r = 0.82$), and internal consistency is moderately high ($\alpha = 0.72$), at least for male respondents. Other self-report measures of type A show similar levels of reliability. Prevalence estimates of type A behavior have varied from 50 to 75% depending on the populations studied and assessment techniques employed.

The validity of self-report measures of type A were originally documented by comparing scores between type A and type B individuals already categorized using the gold-standard interview technique. Although both interview and self-report assessment techniques have good reliability, the two approaches are only modestly interrelated. For example, the JAS showed a 73% agreement with the type A interviews in a cross-validation sample of 419 men participating in the Western Collaborative Group Study. Similarly, the Bortner scale was shown to explain only 53% of the variability in type A clinical interview ratings. This has led to considerable debate as to whether self-report questionnaires of type A are a valid means of describing the same constellation of features identified by the structured interview. Studies using self-report measures have generally provided less association with cardiovascular health outcomes than have interview techniques. As a result, self-report measures of type A have been criticized for relying on the insight of respondents, restricting the range of type A traits, and being influenced by the social undesirability of many items. Because self-report measures do not involve a live interaction with the individual, these scales may not accurately assess the action–emotion complex described by the creators of the interview technique.

Empirical Support

Data on the association between type A and CHD have been inconsistent. Three early large-scale epidemiological studies supported an association between type A and CHD incidence: the Western Collaborative Group in 1975, the Framingham Heart Study in 1980, and the French-Belgian Cooperative Study Group in 1982. All three investigations found that type A individuals had a significantly increased risk of CHD in individuals followed closely for 5–8 years. In the Western Collaborative Group study, 3200 employed men with no heart disease were followed for 8 years. Type A individuals (by structured interview) were found to have a 2.2 relative risk of CHD relative to type B individuals. The Framingham Heart Study, which included a more representative sampling of both men and women, showed relative risk ratios for white-collar men of 2.9 for total CHD, 7.3 for MI, and 1.8 for angina without MI; however, risks were substantially less for women and for blue-collar men. The French-Belgian Cooperative Heart Study followed 2811 male civil servants and factory workers for an average of 5 years. The relative risk ratio for type A men was 1.8 for total CHD. The latter two studies used self-report measures of type A.

Other subsequent epidemiological studies showed little or no relationship. For example, in the Normative Aging Study, correlations of the JAS type A score with CHD risk were not statistically significant. In a 40-year follow-up of 280 men from the 1947 Cardiovascular Disease Project at the University of Minnesota, no statistically significant association was found between the MMPI Type A Scale and CHD. In a large-scale 7-year prospective study (Multiple Risk Factor Intervention Trial, MRFIT), there was no association between type A and risk of coronary events, and this applied to both interview and self-report methods of type A classification. The type A–CHD relation has been much smaller in prospective versus cross-sectional studies, in meta-analyses synthesizing results from multiple studies, and among more diverse groups. Although many plausible explanations have been offered for differences among study results, meta-analyses and other attempts to synthesize results across studies have failed to provide consistent support for type A as a unique and valid predictor of CHD.

The generalizability of initial research linking type A to CHD has also been criticized because early studies were limited to middle-aged, working, White American males. Among more diverse populations, associations between type A and CHD have been weaker or nonexistent. Studies of both type A and hostility in Japan, for example, have shown little association with CHD. Instead, a type C behavior pattern, characterized by a job-centered lifestyle and social dominance at work, was predictive of CHD in working Japanese men. The type A construct may not apply for women due to the greater social unacceptability of type A traits for women.

Intervening biological mechanisms to explain the type A–CHD association have focused primarily on

the hyperreactivity of the cardiovascular and neuroendocrine systems; however, results have been inconsistent. Quantitative reviews have suggested some heightened reactivity in systolic blood pressure response and catecholamine response, but results vary by gender, type A-assessment technique, and the nature of laboratory stressors. Ambulatory blood pressure studies have also indicated some relationship between type A and cardiovascular reactivity. Type A scores have been related to the actual extent of coronary artery atherosclerosis, as judged from cardiac catheterization. Here again, initial studies were positive, but subsequent studies failed to reveal such links.

Other studies have included type A as a potential mediator of health behavior. For example, type A individuals underestimate their degree of physical exertion in response to exercise. They also work longer and at greater intensity when given a self-selected work-pace task. This may explain some of the relationships between type A and cardiovascular health outcomes such as recurrent MI or unexplained cases of sudden death.

Despite the huge accumulated literature on type A, hostility, and CHD, interventions targeting type A behavior are rare. One exception is a 14-month type A counseling approach developed at Mount Zion Hospital purportedly to change belief systems, engage in exercises to modify a sense of time urgency or free-floating hostility, and increase perceptions of personal security. This intervention has been reported to reduce the frequency of silent myocardial ischemia; however, the studies have been performed on small nonrepresentative samples of post-MI coronary patients. It is not known which therapeutic techniques included in this broad-based intervention were most effective.

Component Factors

Because the interview assessment of type A included a number of signs and symptoms, researchers became interested in the specific component factors that might contribute to CHD risk and whether type A was unique in this respect compared with other personality variables. These investigations ultimately led to a shift in the focus of CHD behavioral studies away from the type A global constellation of symptoms. When the total score and subscale scores of the JAS were contrasted, the hard-driving competitiveness factor was more strongly associated with cardiovascular disease than was speed and impatience, job involvement, or the overall type A score. Other studies showed support for hostility as the key underlying component. For example, of 12 behavioral characteristics extracted from the structured interview, a reanalysis of the Western Collaborative Group Study data showed that the hostility component alone accounted for all of the variance explained by type A.

In their 1987 meta-analysis, Booth-Kewley and Friedman identified 18 overlapping personality variables from various assessment measures for the purposes of comparing relative strengths of association. These included global type A scores; subscale measures of speed and impatience, job involvement, hard-driving competitiveness, and time urgency; and non-type A measures of anger, hostility, aggression, depression, extraversion, and anxiety. These analyses showed that depression was more strongly associated with CHD than was type A, and CHD correlations with hostility, aggression, and extraversion were of a similar magnitude to type A. This meta-analysis and subsequent studies ultimately led to the conclusion that only the type A components related to anger, hostility, and aggression were responsible for increased CHD risk. In particular, hostility, which had long been suspected as a risk factor for CHD, was found to explain nearly all the shared variance between type A and CHD.

Today, the expanding interest in hostility as a predictor of CHD has eclipsed the original global type A as a potential risk factor. However, conceptual and assessment controversies continue as various components of hostility are hypothesized to underlie associations with CHD. Cynical hostility, as measured by the Cook–Medley scale of the MMPI, was shown to predict a fivefold increase in CHD risk in a 25-year follow-up of 255 medical students. Suppressed anger was associated with a twofold increase in mortality among 696 Michigan residents followed over 12 years. Studies linking hostility to CHD have shown associations with angiographically documented CHD. However, their cross-sectional nature precludes definitive conclusions regarding cause and effect. Prospective studies with 5- to 25-year follow-ups have shown associations between hostility and later CHD development, but similar associations exist with all-cause mortality. Therefore, the specificity of hostility with regard to CHD remains unclear.

Beyond Coronary Heart Disease

Although the focus of type A research has been primarily on CHD, hypertension, and related cardiovascular health outcomes, several studies have examined the impacts of type A behaviors in the context of entirely different medical conditions. These have included chronic fatigue syndrome, irritable bowel syndrome, and other health problems thought to involve

psychological risk factors. These studies have generated no consistent pattern of findings that implicate type A as a central causal factor, although it periodically emerges as a significant predictor of various health outcomes. The most common explanation for this finding is that type A behavior may play a mediating role in health behaviors (exercise, diet, smoking, etc.).

Despite the popular use of the type A label to describe the overachieving efforts of co-workers and family members, there have been relatively few studies of the interpersonal and occupational consequences of type A behavior. In cross-sectional studies, type A overlaps with other measures of personality and temperament (e.g., narcissism, optimism, and aggression), yet type A has not emerged as an independent mental health risk factor. In relation to perceptions of stress and satisfaction at work and home, the impacts of type A behavior have been inconsistent. Based on the available evidence, the type A constellation of personality traits is an uncertain predictor of social and emotional well-being.

Conclusion

Any fair-minded observer would be immediately impressed by the numerous links between behavior and coronary disease. The impact of various risk-enhancing and risk-reducing behaviors is unchallenged. Similarly, the impact of unusual life stressors in provoking a deterioration in already-existing CHD is reasonably established. It is less clear that a habitual personality or behavior pattern leads to the development of CHD. Ambiguities in recognizing that personality and defining its central features have plagued this area. Nonetheless, data do suggest some links between type A behavior and CHD. Current work focuses less on type A itself than on the quest for its toxic core, its mechanism of action, and the possibilities of intervention.

See Also the Following Articles

Aggression; Anger; Blood Pressure; Cardiovascular System and Stress; Heart Disease/Attack; Hostility; Metabolic Syndrome; Self-Esteem, Stress and Emotion.

Further Reading

Booth-Kewley, S. and Friedman, H. (1987). Psychological predictors of heart disease: a quantitative review. *Psychological Bulletin* **101**, 343–362.

Dimsdale, J. (1988). A perspective on type A behavior and coronary disease. *New England Journal of Medicine* **318**, 110–112.

Friedman, M. and Ghandour, G. (1993). Medical diagnosis of type A behavior. *American Heart Journal* **126**, 607–618.

Friedman, M. and Powell, L. H. (1984). The diagnosis and quantitative assessment of type A behavior: introduction and description of the videotaped structured interview. *Integrated Psychiatry* **2**, 123–129.

Friedman, M. and Rosenman, R. (1974). *Type A behavior and your heart*. New York: Knopf.

Gallacher, J. E. J., Sweetnam, P. M., Yarnell, J. W. G., et al. (2003). Is type A behavior really a trigger for coronary heart disease events? *Psychosomatic Medicine* **65**, 339–346.

Grunbaum, J., Vernon, S. W. and Clasen, C. M. (1997). The association between anger and hostility and risk factors for coronary heart disease in children and adolescents: a review. *Annals of Behavioral Medicine* **19**, 179–189.

Harbin, T. J. (1989). The relationship between the type A behavior pattern and physiological responsivity: a quantitative review. *Psychophysiology* **26**, 110–119.

Haynes, S. G., Feinleib, M. and Kannel, W. B. (1980). The relationship of psychosocial factors to coronary heart disease in the Framingham Study. III: Eight-year incidence of coronary heart disease. *American Journal of Epidemiology* **111**, 37–58.

King, K. B. (1997). Psychologic and social aspects of cardiovascular disease. *Annals of Behavioral Medicine* **19**, 264–270.

Matthews, K. and Hanes, S. (1986). The type A behavior patterns and coronary disease risk: update and critical evaluation. *Journal of Epidemiology* **123**, 923–960.

Melamed, S., Harari, G. and Green, M. S. (1993). Type A behavior, tension, and ambulatory cardiovascular reactivity in workers exposed to noise stress. *Psychosomatic Medicine* **55**, 185–192.

Review Panel on Coronary-Prone Behavior and Coronary Heart Disease (1981). Coronary prone behavior and coronary heart disease: a critical review. *Circulation* **63**, 1199–1215.

Thoreson, C. E. and Powell, L. H. (1992). Type A behavior pattern: new perspectives on theory, assessment, and intervention. *Journal of Consulting & Clinical Psychology* **60**, 595–604.

U

Ubiquination *See:* Proteosome; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Ulceration, Gastric

R Murison and A M Milde
University of Bergen, Bergen, Norway

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Murison, volume 3, pp 631–633, © 2000, Elsevier Inc.

Background
Psychological Influences on Development of Ulceration
Mechanisms
Future Directions

Glossary

Analgesia	Insensibility to painful stimuli.
Corticotropin releasing hormone	A hormone released from the hypothalamus which leads to secretion of cortisol from the adrenal cortex into the circulation.
Dopamine	A catecholamine neurotransmitter.
Gastric acid	Hydrochloric acid secreted by cells in the stomach.
Helicobacter pylori (*H. pylori*)	A gram-negative bacterium found in the majority of patients with peptic ulcer.
Immunomodulation	Changes in the function of the body's immune defence system.
Ischemia	Localized oxygen deficit as a result of obstruction of the blood supply or vasoconstriction.
Lactobacillus	A bacterium found particularly in dairy produce and meat.
Limbic	Referring to several phylogenetically old brain structures thought to be central to emotionality and learning.
Mucosa	The mucous membrane lining the stomach and intestines.
Pepsinogen	The precursor of pepsin, found in the stomach mucosa.
Peptic ulcer	An ulcer involving the mucosa, submucosa, and muscular layers of the stomach, or duodenum, due in part at least to the action of gastric juices.
Rumen	The part of the stomach lacking acid-producing glands.
Spontaneously hypertensive rats	A specially bred strain of rats with high susceptibility to develop high blood pressure.
Sympathetic	The part of the autonomic nervous system which prepares the organism for the fight/flight response.
Vagotomy	Severing of the vagus nerve.
Vagus nerve	The tenth cranial nerve forming an important part of the parasympathetic system.

Background

Ulcerations of the stomach, typically manifested as nonpenetrating erosions of the gastric mucosa, are found following stress in many mammalian species (mice, rats, horses, primates, swine, reindeer, and dogs, among others). The condition is therefore considered to be a reflection of fundamental common psychobiological processes activated by stress. Although there are significant exceptions, most psychobiological studies have focused on the rat because of the extensive pool of knowledge of behavior and central nervous system (CNS) function in that species. The ulcerations are found in both the rumen and the glandular areas of the stomach. The former appear to be the result of food deprivation procedures, whereas the latter are sensitive to psychological factors (stress). The ulcerations are typically bleeding erosions of the superficial gastric mucosa along the

ridges of the gastric folds. Prolonged application of the stressor may cause full-blown ulcers, with the erosions reaching through the muscularis mucosa. Given a respite from the stressor and other potential ulcerogenic factors, however, the ulcerations typically heal within a few days.

Controversy surrounds the homologous condition in humans. Earlier, gastric ulcerations in rodents and monkeys were used as models of peptic ulcer in humans, even though the location of ulcer in humans is more typically the duodenum. Later, gastric ulcerations were regarded as homologous to the acute stress ulcers seen after burn injuries or traumatic surgery in humans. Since the discovery of the role of *Helicobacter pylori* in peptic ulcers in humans, it has been surmised that ulcerations in animals are homologous to the erosive prepyloric changes of the gastric mucosa reported in subgroups of patients diagnosed with functional dyspepsia. However, there is much more to the story.

Psychological Influences on Development of Ulceration

In the wild, ulcerations are observed in animals (both rodents and primates) following social defeat. In both wild and domestic animals, ulcerations have been reported in connection with the stress of transport and confinement. In the laboratory, ulcerations are produced by a wide variety of stressors in animals – electric shock, restraint, cold, activity stress, social stress, or combinations of these – with the purpose of exploring the role of psychological factors that may modulate the incidence and/or severity of the mucosal injury. Consciousness is a modulating, if not necessary, condition for the development of ulcerations.

Ulcerations played a key role in Selye's description of the general adaptation syndrome and constituted one of the elements of this nonspecific response to diverse nocuous agents, together with shrinkage of the thymus and enlargement of the adrenal glands.

Acute Effects

To identify the psychological factors in ulceration (and other stress-related phenomena), many have employed the triadic design. Animals are tested in triplets. One animal (the executive) is subjected to a physical stressor (most commonly electric shock) that can be escaped from or avoided by instrumental responses. The second (yoked) animal is wired to the first such that the same physical stress is experienced but there is no possibility of avoidance or escape. The third (control) animal is subjected to the same apparatus conditions as the first two but not to the stressor. By comparing the amount of ulceration between the first two animals, the psychological and physical components of the stressor can be identified.

The Executive Monkey studies with monkeys, by Brady and his colleagues, were based on a triadic design, and indicated that the executive was more vulnerable to ulceration than its yoked partner. These results were widely accepted at the time, but their interpretation was later questioned. In contrast to the Brady results, Weiss reported that executive rats developed less ulceration than their yoked partners. These inconsistencies between the monkey and rat results have been attributed to the particular group-assignment procedures and experimental parameters used by Brady rather than to species differences.

Later, Weiss demonstrated that ulceration in rats was attenuated by providing the animals with predictability for shock and was exacerbated by either increasing the response requirements (effort) required to maintain instrumental control or by introducing a component of psychological conflict. Weiss's model predicts ulceration as a function of the effort required to maintain instrumental control over a stressor and an inverse function of the feedback provided by such control. Successful coping behaviors (high feedback) ameliorate ulceration, but greater workload (effort) will increase ulceration. This model has played an important role in the development of recent theories concerning the relationships among stress, coping, and ill health in humans. Between- and within-strain studies have led to the notion that more emotional animals are more vulnerable to ulcerations than less emotional conspecifics. For example, spontaneously hypertensive rats are less vulnerable than the more emotional Wistar–Kyoto strain. Potential biological markers for vulnerability to ulceration include plasma levels of pepsinogen (also in humans) and dopaminergic function. Genetic factors certainly play a role in vulnerability to ulceration, but these are modulated by early life experiences and previous experiences of coping/noncoping, which also modulate dopaminergic function.

Proactive Effects

A number of experimenters have investigated proactive effects of one or more stressor exposures on vulnerability to ulceration under later exposure to either the same or a different stressor. Rats previously subjected to the stress of whole-body restraint generally develop less ulceration under later restraint, whereas previous exposure to an activity wheel primes the animal for later activity-stress ulceration. Using heterogeneous stress experiences, rats subjected to the same shock stress parameters that are used to induce learned helplessness (a putative model for

human depression) develop greater amounts of ulceration under a later radically different type of stressor, restraint in water. This phenomenon is mediated by opioid mechanisms in the same way that stress-induced analgesia and stress-induced immunomodulation are also invoked by these same shock parameters. Previous experience with escapable shock, on the other hand, has proactive protective effects against later ulceration, although the mechanisms of this phenomenon remain unexplored. More recent studies have complemented these earlier reports in showing that prior experience with seminaturalistic coping responses is protective against later ulcerogenesis. Similarly, prior experience with uncontrollable shocks, but in which safety signals are provided at the end of each shock, also reduces vulnerability under later stress.

Early weaning of rats (at day 16) leads to a drastic increase in vulnerability to ulceration during a limited period later on (22–40 days). Although later studies indicated that this effect was due more to nutritional and thermoregulatory influences than to psychobiological phenomena, recent reports suggest that such very early manipulations have profound long-term effects on dopaminergic system function, which is known to play an important role in ulceration.

Mechanisms

Peripheral Mechanisms

The peripheral mechanisms of ulceration were originally believed to be limited to increased levels of vagally mediated gastric acid secretion, and vagotomy does reduce ulceration. Although the old saying "no acid – no ulcer" still holds, recent approaches have focused on combinations of changes in gastric motility, gastric blood flow, and sympathetic tone. A common view is that enhanced vagally mediated acid secretion and reduced gastric blood flow represent key aggressive factors, whereas the protective factors include prostaglandin-mediated secretion of mucous, secretion of bicarbonate, and maintenance of an adequate blood supply.

The development of ulceration appears to be greatest in the period immediately following an acute stress. This has led to the concept of ulceration as a poststress or rebound phenomenon, whereby gastric acid secretion is inhibited during the application of a strong stressor (although certain stressors may increase secretion). A concurrent strong sympathetically mediated constriction of the mucosal blood vessels increases the risk for local ischemia, and slow rhythmic contractions may cause further physical damage. On release from the inhibitory (sympathetic) influence of the acute stressor, gastric acid levels are rapidly elevated, leaving a vulnerable mucosa open to damage.

Interactions between stress and bacterial flora The role of bacterial flora is now under intensive study since the discovery that ulcer disease in humans is related to the presence of *H. pylori*. Animal psychobiological models for *Helicobacter* infection have been hindered by the fact that most behavioral studies of ulceration have been performed in rats, whereas the models for infection have largely been limited to mice, ferrets, and gerbils. It remains unclear whether bacterial flora play any part in either ulceration in animals or erosive prepyloric erosions in humans. Germ-free animals do develop stress gastric ulcerations, although to a lesser extent than normal animals, and antibiotic treatment accelerates the healing of the ulceration. From the human data, it is apparent that factors additional to *H. pylori* are involved in gastric pathology, and it remains to be clarified to what extent stress, either alone or in combination with bacterial factors, is involved. Recent results also suggest that the relationship between the bacterial population of gut and ulceration is two way: The induction of ulceration in animals by the direct application of acetic acid to the stomach lining dramatically changes the predominant bacterial species, which then return to normal as the ulceration heals. In particular, *Lactobacillus* seems to be protective and enhances the healing process. Consistent with these animal data, a small-scale study in humans has shown that *H. pylori*, although present in all patients with duodenal ulceration with a duration greater than 6 months, was not present in patients with a shorter history, suggesting that the infection is not so much a cause as a factor producing chronicity. On the other hand, a recent study of patients admitted to intensive care units shows that *H. pylori* infection was more frequent in patients exhibiting upper gastrointestinal bleeding than in patients who were not bleeding. Thus, the controversy concerning the role of *H. pylori* remains unresolved.

Central and Neuroendocrine Influences

Studies of the central mechanisms of ulcerations have focused on limbic structures, in particular the amygdala (with its close association with the vagus nerve), the septum, and the hippocampus. The stimulation of the amygdala elicits ulceration, whereas lesions of the central nucleus of the amygdala attenuate their formation. The stimulation of the dorsolateral septum reduces the amount of ulceration induced by cold-restraint stress. Lesions of the hippocampus, in particular the ventral hippocampus, exacerbate

ulceration, presumably acting through the amygdala-vagal pathways.

Studies of centrally acting peptides show that corticotropin releasing hormone (CRH) has a potent protective effect when applied directly in the central amygdala. When injected peripherally, CRH protects against stress ulceration in young, but not old, rats (in old rats, the effect is exacerbatory). CRH acts through sympathetic pathways to inhibit gastric acid secretion while at the same time stimulating bicarbonate production and inhibiting slow gastric contractions. Tests of the CRH type-1 receptor antagonist antalarmin demonstrate a potent anti-ulcerogenic effect, but the glucocorticoid/progesterone antagonist RU-38486 significantly potentiates the formation of stress-induced gastric erosions. The role of the CRH system in ulcerogenesis remains unclear and peripheral corticosterone may play an important role in cytoprotection during stress.

Thyrotropin releasing hormone (TRH) acts as a potent ulcerogenic agent and is stimulated by cold. TRH promotes both gastric acid secretion and the slow gastric contractions associated with the development of ulcerations induced by cold restraint stress. The main site of action for TRH appears to be the dorsal motor nucleus of the vagus, although a number of hypothalamic nuclei and the amygdala are also implicated.

Adrenalectomy increases the incidence of ulceration, and there is little support for the earlier notion that corticosterone plays any causal role in ulcerogenesis. Rather, the events likely to activate the pathogenic mechanisms for ulceration are also the kind of events that activate the hypothalamic-pituitary-adrenal axis.

Of the centrally acting neurotransmitters, most attention has been paid to dopamine and in particular to the D1 and D2 receptors. In line with the observation that schizophrenics rarely develop ulcers, whereas untreated Parkinson's patients are at high risk, numerous studies have provided evidence that central dopamine is protective against ulceration. A more complex picture has now become apparent because results suggest that the stimulation of dopamine D1 receptors inhibits ulceration whereas stimulation of dopamine D2 receptors has a pro-ulcerogenic effect. Within-strain comparisons of rats rated as high emotional versus low emotional reveal that the former are more vulnerable to stress ulcerations than the latter, as well as having lower levels of dopamine in the amygdala. A similar difference between these two types of animal was observed with respect to ethanol-induced ulcerations. Rats differing in their sensitivity to the dopaminergic agonist apomorphine are differentially susceptible to ulcerations and recover at different rates, and experience with shock parameters demonstrated to increase vulnerability to ulceration change their apomorphine sensitivity accordingly. Recent evidence also suggests that the dopamine content of the gastric mucosa is important in cell proliferation of the gastric mucosa. These and other data suggest that vulnerability profiles of animals to ulcerations are manifest at both the central and peripheral levels.

Future Directions

Studies on the etiology of stress gastric ulceration have become reduced in number since the experience that *H. pylori* eradication is the most effective treatment for human peptic ulcer. However, the bacterial risk factor and the effectiveness of treatment do not in themselves account for either historical or contemporary links between stress and ulcerogenesis. Recent developments in psychoneuroimmunology offer opportunities to study how host factors (including CNS-mediated immunomodulation) may interact with pathophysiological agents, both bacterial and others, in causing and maintaining gastric mucosal changes. Specifically, stress-related failures of immune defenses may provide a window of opportunity for *H. pylori* to cause problems.

See Also the Following Article

Gastrointestinal Effects.

Further Reading

Boulos, P. B., Botha, A., Hobsley, M., et al. (2002). Possible absence of *Helicobacter pylori* in the early stages of duodenal ulceration. *Quarterly Journal of Medicine* **95**, 749–752.

Degen, S. B., Geven, E. J., Sluyter, F., et al. (2003). Apomorphine-susceptible and apomorphine-unsusceptible Wistar rats differ in their recovery from stress-induced ulcers. *Life Sciences* **72**, 1117–1124.

Desai, J. K., Goyal, R. K. and Parmar, N. S. (1999). Characterization of dopamine receptor subtypes involved in experimentally induced gastric and duodenal ulcers in rats. *Journal of Pharmacy and Pharmacology* **51**, 187–192.

Elliott, S. N., Buret, A., McKnight, W., et al. (1998). Bacteria rapidly colonize and modulate healing of gastric ulcers in rats. *American Journal of Physiology* **275**, G425–G432.

Filaretova, L., Podvigina, T., Bagaeva, T., et al. (2001). Gastroprotective action of glucocorticoids during the formation and the healing of indomethacin-induced gastric erosions in rats. *Journal of Physiology* (Paris) **95**, 201–208.

Gabry, K. E., Chrousos, G. P., Rice, K. C., et al. (2002). Marked suppression of gastric ulcerogenesis and intestinal responses to stress by a novel class of drugs. *Molecular Psychiatry* **7**, 433, 474–483.
Glavin, G. B., Murison, R., Overmier, J. B., et al. (1991). The neurobiology of stress ulcers. *Brain Research Reviews* **16**, 301–343.
Levenstein, S. (2000). The very model of a modern etiology: a biopsychosocial view of peptic ulcer. *Psychosomatic Medicine* **62**, 176–185.
Maury, E., Tankovic, J., Ebel, A., et al. (2005). An observational study of upper gastrointestinal bleeding in intensive care units: is *Helicobacter pylori* the culprit? *Critical Care Medicine* **33**, 1513–1518.
Murison, R. (2001). Is there a role for psychology in ulcer disease? *Integrative Physiological and Behavioral Science* **36**, 75–83.
Overmier, J. B. and Murison, R. (2000). Anxiety and helplessness in the face of stress predisposes, precipitates, and sustains gastric ulceration. *Behavioral Brain Research* **110**, 161–174.
Sapolsky, R. M. (2004). *Why zebras don't get ulcers* (3rd edn.). New York: Henry Holt.
Szabo, S. (1979). Dopamine disorder in duodenal ulceration. *Lancet* **2**(8148), 880–882.
Weiner, H. (1991). From simplicity to complexity (1950–1990): the case of peptic ulceration. I: Human studies. *Psychosomatic Medicine* **53**, 467–490.
Weiner, H. (1991). From simplicity to complexity (1950–1990): the case of peptic ulceration. II: Animal studies. *Psychosomatic Medicine* **53**, 491–516.

Ultradian Rhythms

S L Lightman
University of Bristol, Bristol, UK

© 2007 Elsevier Inc. All rights reserved.

Glossary

Adrenocorticotropic hormone (ACTH)	A peptide secreted by the anterior pituitary gland that stimulates the secretion of adrenal glucocorticoids; also called corticotropin.
Arginine vasopressin (AVP)	A hypothalamic peptide that acts at the kidney to moderate water excretion (hence, also termed the antidiuretic hormone) and acts synergistically with CRH to stimulate ACTH release from the anterior pituitary gland.
Corticotropin releasing hormone (CRH)	A brain peptide that mediates the neural control of ACTH release from the anterior pituitary gland and that has also been implicated in central neural actions related to anxiety and depression.
Follicle-stimulating hormone (FSH)	An anterior pituitary gonadotropin responsible for stimulating the growth of ovarian follicles and the secretion of estrogen in the female and the development of sperm in the male.
Glucocorticoid receptor (GR)	A receptor for glucocorticoids, of which cortisol is the main type in the human.
Gonadotropin-releasing hormone (GnRH)	A peptide that is secreted by the hypothalamus and that mediates the neural control of pituitary gonadotropin secretion.
Growth hormone (GH)	An anterior pituitary hormone that controls body growth and metabolism.
Hypothalamic-pituitary-adrenal (HPA) axis	One of the two main neuroendocrine stress response systems.
Janus kinase (JAK)	A kinase with two phosphorylation sites, only one of which appears to be active.
Luteinizing hormone (LH)	An anterior pituitary peptide that induces ovulation and the secretion of progesterone in the female and the secretion of testosterone in the male.
Mineralocorticoid receptor (MR)	A receptor for mineralocorticoid steroid hormones, of which aldosterone is the prototype.
mRNA	A molecule that transcribes the genetic code from DNA.
Signal transducer and activator of transcription 5 (STAT5)	A transcription factor implicated in cell differentiation and tumor formation.

Frequency encoding of intercellular signals is a well-accepted mode of communication between neurons. More than this, however, it is actually a common mechanism of communication across a broad range of both inter- and even intracellular moieties. Even an organism as primitive as the slime mold (*Dictyostelium discoideum*) aggregates only in response to external pulses of cAMP delivered with a periodicity

of 5 min and not to constant stimuli or frequencies greater than every 2 min.

Within mammalian systems, frequency encoding is mediated by circulating hormones as major signaling molecules, and many of these signal through the mechanisms of ultradian rhythmicity. Thus, pulsatile GnRH release is essential for maintenance of LH and FSH secretion, and the modulation of GnRH pulse frequency can produce differential LH and FSH secretion via regulation of LHβ and FSHβ mRNA expression. Similarly, sex differences in episodic release of growth hormone (GH) elicit significant differences in gene expression. High-amplitude GH bursts occur at intervals of 3–4 h in male rats, whereas more frequent low-amplitude pulses occur in females; these sexually diergic patterns are responsible for marked differences in the expression of liver enzymes and JAK-dependent STAT5 responses.

Recent advances in our understanding of glucocorticoid receptor function suggest potential mechanisms for similar regulation of gene expression in cells exposed to pulsatile changes in glucocorticoid concentration. At the level of the glucocorticoid receptor itself, Dekelbab et al described in 2004 a rapid downregulation of GR associated with prolonged exposure to glucocorticoids that can be avoided if corticosterone is administered intermittently. This suggests a more effective induction of GR target genes from intermittent pulses of glucocorticoids. Furthermore, it is now recognized that the binding of the GR to the promoter regions of target genes is not the static process that was previously accepted. The availability of GR tagged with fluorescent tags and the advent of chromatin immunoprecipitation (ChIP) techniques have revealed that the interaction among the GR, DNA, and other proteins involved in transcription is a highly dynamic process. Indeed the effective concentration of nuclear receptors on some promoters appears to vary after hormone stimulation in a wavelike or cycling fashion. Métivier et al. described a transcriptional clock with another steroid ligand receptor – the estrogen receptor – engaging in 40-min cyclic occupation of a target promoter.

The HPA axis shows classic hormone rhythmicity. In the basal state, normal HPA activity is characterized both by a circadian rhythm of cortisol and corticosterone and also an ultradian rhythm that results in a pulsatile release of ACTH and glucocorticoids throughout the 24-h cycle. This ultradian rhythmicity is clearly a fundamental feature of HPA activity and *ex vivo* evidence from the monkey suggests that the pulse generator is located within the hypothalamus. This pulse generator results in the phasic release of CRH and AVP from the parvocellular neurons of the hypothalamic paraventricular nucleus into the portal blood, from which it has access to pituitary corticotrophs that secrete ACTH, which in turn regulates the production and release of glucocorticoids from the fasciculata cells of the adrenal cortex.

Details of studies on the intrinsic dynamics of HPA function in free-running conscious animals reveals a complex ultradian rhythm of corticosterone secretion of approximately 1 h in duration. This profile shows not only marked sexual diergism, but also shows marked plasticity in different physiological states such as lactation and aging, the presence of or susceptibility to disease, and early life programming. Similar ultradian rhythmicity is found in humans, resulting in complex patterns of cortisol secretion that have made the response to stress testing difficult to interpret.

The rapidly changing concentrations of glucocorticoids have major potential implications with respect to the pharmokinetics of receptor occupancy and signal transduction because, over relatively short periods of time, there are widely changing concentrations of ligand available for both MR and GR binding. Although there is little data on this, it appears likely that this is of a special relevance for GR occupancy because the high affinity of the MR suggests that short-term changes in ligand concentration have only relatively small effects on MR signal transduction.

Pulsatility provides an added dimension to the regulation of hormone responsive pathways. Not only can it provide an additional mode of signaling within the same amount of total hormone secreted, but it could also potentially have differential effects on different tissues within the same organism. This could be picked up both by different patterns of expression of the hormone receptors (e.g., MR and GR) and by differential effects on receptor recycling, postreceptor signal transduction, or interaction with other intracellular pathways.

See Also the Following Articles

Circadian Rhythms, Effects of Prenatal Stress in Rodents; Circadian Rhythms, Genetics of; Seasonal Changes in Stress Responses; Sleep Loss, Jet Lag, and Shift Work; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress; Cortisol Awakening Response; Night Shiftwork; Seasonal Rhythms; Circadian Rhythm Effects on Cardiovascular and Other Stress-Related Events.

Further Reading

Belchetz, P. E., Plant, T. M., Nakai, Y., et al. (1978). Hypophysial responses to continuous and intermittent delivery of hypothalamic gonadotropin-releasing hormone. *Science* **202**, 631–633.

Darmon, M., Brachet, P. and Da Silva, L. H. (1975). Chemotactic signals induce cell differentiation in *Dictostelium discoideum*. *Proceedings of the National Academy of Sciences USA* **72**, 3163–3166.

Dekelbab, B. H., Witchel, S. F. and De Franco, D. B. (2004). Continuous versus intermittent administration of glucocorticoids: ligand-dependent down-regulation of glucocorticoid receptors revisited. American Endocrine Society meeting, abstract P2–131.

Folenius, M., Simon, C., Brandenberger, G., et al. (1987). Ultradian plasma corticotropin and cortisol rhythms: time-series analyses. *Journal of Investigative Endocrinology* **10**(3), 261–266.

Goldbeter, A. (1996). *Biochemical oscillations and cellular rhythms: the molecular basis of periodic and chaotic behaviour.* Cambridge UK: Cambridge University Press.

Iranmanesh, A., Lizarralde, G. and Veldhuis, J. D. (1993). Co-ordinate activation of the corticotropic axis by insulin-induced hypoglycemia: simultaneous estimates of beta-endorphin, adrenocorticotropin and cortisol secretion and disappearance in normal men. *Acta Endocrinologia (Copenhagen)* **128**(6), 521–528.

Lightman, S. L., Windle, R. J., Julian, M. D., et al. (1999). Significance of pulsatility in the HPA axis. In: Chadwick, D. J. & Goode, J. A. (eds.) *Symposium on mechanisms and biological significance of pulsatile hormone secretion, Novartis Foundation Symposium 277,* Chichester, UK: John Wiley.

Mershon, J. L., Sehlhorst, E. S., Rebar, R. W., et al. (1992). Evidence of a corticotropin-releasing hormone pulse generator in the macaque hypothalamus. *Endocrinology* **130**, 2991–2996.

Métivier, R., Penot, G., Hübner, M. R., et al. (2003). Estrogen receptor-α dissects ordered, cyclical and combinatorial recruitment of cofactors on a natural target promoter. *Cell* **115**, 751–763.

Nolan, L. A., Windle, R. J., Wood, S. A., et al. (2000). Chronic iodine deprivation attenuates stress-induced and diurnal variation in corticosterone secretion in female Wistar rats. *Journal of Neuroendocrinology* **12**, 1149–1159.

Papavasiliou, S. S., Zmeili, S., Khoury, S., et al. (1986). Gonadotropin-releasing hormone differentially regulated expression of the genes for luteinizing hormone α and β subunits in male rats. *Proceedings of the National Academy of Sciences USA* **83**, 4026–4029.

Seale, J. V., Wood, S. A. and Atkinson, H. C. (2004). Gonadectomy reverses the sexually diergic patterns of circadian and stress-induced hypothalamic-pituitary-adrenal axis activity in male and female rats. *Journal of Neuroendocrinology* **16**, 516–524.

Shanks, N., Windle, R. J., Perks, P. A., et al. (2000). Early-life exposure to endotoxin alters hypothalamic-pituitary-adrenal function and predisposition to inflammation. *Proceedings of the National Academy of Sciences USA* **97**(10), 5645–5650.

Waxman, D. J., Ram, P. A., Park, S. H., et al. (1995). Intermittent plasma growth hormone triggers tyrosine phosphorylation and nuclear translocation of a liver-expressed, stat 5-related DNA binding protein: proposed role as an intracellular regulator of male-specific liver gene transcription. *Journal of Biological Chemistry* **270**, 13262–13270.

Wildt, L., Hausler, A., Marshall, G., et al. (1981). Frequency and amplitude of gonadotropin-releasing hormone stimulation and gonadotropin secretion in the rhesus monkey. *Endocrinology* **109**(2), 376–385.

Windle, R. J., Wood, S. A., Kershaw, Y. M., et al. (2001). Increased corticosterone pulse frequency during adjuvant-induced arthritis and its relationship to alterations in stress responsiveness. *Journal of Neuroendocrinology* **13**, 905–911.

Windle, R. J., Wood, S. A., Lightman, S. L., et al. (1998). The pulsatile characteristics of hypothalamo-pituitary-adrenal activity in female Lewis and Fischer 344 rats and its relationship to differential stress responses. *Endocrinology* **139**, 4044–4052.

Windle, R. J., Wood, S. A., Shanks, N. M., et al. (1998). Ultradian rhythm of basal corticosterone in the female rat: dynamic interaction with the response to acute stress. *Endocrinology* **139**, 443–450.

Understimulation/Boredom

V J Sutherland
Sutherland-Bradley Associates, Staffordshire Moorlands, UK

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by V J Sutherland, volume 3, pp 634–636, © 2000, Elsevier Inc.

Understimulation and Boredom as a Rustout Syndrome
Understimulation and the Stress Response
Conclusion

Glossary

Boredom	The state of being tired or made weary by being dull, repetitious, or uninteresting; an emotion in response to an environment that is perceived to be monotonous and that fails to stimulate the individual.
Understimulation	Insufficient arousal in respect to the senses or activity in the individual.

Understimulation and Boredom as a Rustout Syndrome

While a plethora of research evidence exists to suggest that overstimulation in the form of work overload is a key source of stress, associated with a wide variety of negative health and performance variables, the issue and impact of understimulation and boredom should not be ignored or underestimated. Overload as a form of stress can lead to burnout, whereas the lesser used term of rustout might be useful to explain the negative impact of understimulation that results in boredom, apathy, low morale, and absence from the workplace. It is suggested that an unvaried work environment lacks stimulation, and repeated performance of a routine task can be perceived as monotonous. For example, Persson and colleagues observed that understimulation was a potential problem associated with operator work in automated process control, where monitoring was a predominant part of the job. Such conditions can lead to rustout. Similarly, a medical practitioner might experience rustout when the individual feels that his or her skills and abilities are not being utilized to his or her satisfaction. In both instances, the workplace or job produces feelings or emotions that are described as "boring". When an employee feels able to produce a report but is not permitted to do so because the boss always expects to take over completion of the final document, qualitative underload is experienced. This can also be viewed as understimulation, and the perceived state of boredom or job dissatisfaction is described as the stressor outcomes or stress response. Lack of a stimulating work environment or challenge may be classified as the stressor or source of stress. These form a stress chain, namely, the stressor leads to perception of a state of stress, which leads to the stress response, or coping. However, it must be acknowledged that stress chains and cause-and-effect relationships are not simple or necessarily linear. Boredom, like pain, can not be seen or measured in an objective way because it is a subjective experience, mediated by individual differences such as personality, needs, wants, and past history. While agreement is not unanimous, it is usually acknowledged that there is an association between boredom and a state of low arousal. Grandjean placed boredom in his categories of fatigue, describing it as mental fatigue wherein the individual reports a feeling of weariness. Research evidence suggests that mentally fatigued workers are likely to behave in an inconsistent manner and/or take short cuts, thereby jeopardizing safety and performance standards. Kass and colleagues demonstrated the business costs associated with boredom at work. Workers in a manufacturing plant scoring high on both state and trait boredom were significantly more dissatisfied with the work itself, pay, promotion, supervisors, and co-workers. Those scoring high in job boredom were also more likely to be job absent, but had longer organizational tenure.

Understimulation and the Stress Response

Early research findings presented by Yerkes and Dodson explained that a certain amount of stimulation (i.e., arousal) was needed for optimal performance but that both understimulation and overstimulation could produce negative effects upon the individual in terms of poor health and ineffectiveness. Thus, the curvilinear relationship observed between level of arousal and performance outcome, known as the Yerkes–Dodson Law or the inverted-U hypothesis, clearly illustrates how too little stimulation can lead to boredom, apathy, and ultimately, depression (i.e., rustout), while overstimulation can result in high anxiety, tension, and burnout. However, response to stimulation will vary among individuals because it is a subjective experience

contingent upon one's perceptions of the experience. A tale of stonecutters in the Middle Ages provides an illustration of the role played by our perceptions in a potentially understimulating situation in the workplace. A traveler approached a group of stonecutters and asked, "What are you doing?" The first responded, "I'm cutting stone. It's dull work but it pays the bills." The seconded stonecutter said, "I'm the best stonecutter in the land. Look at the smoothness of this stone, how perfect the edges are." The third man pointed to a foundation several yards away and said, "I'm building a cathedral." Thus, a state of stress is said to exist when an individual perceives an imbalance between the perceived level of demand (i.e., stimulus arousal) and his or her perceived ability to meet that demand.

When the task structure or design of the job is perceived as monotonous or repetitive, it can be demotivating and a barrier to performance effectiveness. Typically, such jobs offer little scope for decision making, autonomy to influence working methods, or pace and sequence of task. The passive job described by Karasek is typically a low-demand, low-autonomy job, and as such will generally be dissatisfying as it induces a decline in overall activity and a reduction in general problem-solving activity. Indeed, Zavala and colleagues observed that workers with a low workload were late to work or left early more often than individuals with an average or high workload. However, not all research findings support the notion that industrial monotony must have negative health or poor productivity effects because expectation determines how the situation is experienced. Hence, the sameness of the job is regarded as normal, and while it might not be liked, there seems to be no expressed job dissatisfaction or harmful health effects. The job is simply viewed as an instrumental means to the ends of obtaining what the employee needs outside of the workplace. The notion of understimulation as lack of or low arousal was expanded by the work of French and Caplan, who explained the concept in terms of workload while distinguishing between quantitative and qualitative load.

Quantitative versus Qualitative Underload

It is possible to experience quantitative underload as understimulation because one simply does not have enough to do to fill the time. Cooper and Kelly reported that crane operators confined to their cab and presented with a very low level of work activity experienced quantitative underload, which was a significant predictor of anxiety, depression, and job dissatisfaction among this group of workers. Boredom in the daily routine as a result of too little to do might result in inattentiveness, which is potentially hazardous, for example, when an employee fails to respond to an emergency situation. Dangers become exacerbated during night-shift work when an individual must adjust to a new sleep pattern but does not have enough to do to stay alert. Cheliout et al. found a high incidence of deactivation episodes among electronics assemblers who were monitored by continuous electroencephalography over a 1-day period. The theta rhythms observed, described as microsleep (i.e., asleep with one's eyes open), were indicative of the boredom and tedium experienced by these workers. When a task does not provide opportunities to use or develop skills and abilities, the situation is described as qualitative underload, and the individual can feel undervalued and bored. This state of apathy can lead to poor morale, irritation, depression, job dissatisfaction, and absenteeism. Thwarted ambition, unfulfilled goals, and unmet expectations cause frustrations that can generate various forms of destructive behavior. Examples of these maladaptive coping strategies include workplace vandalism, sabotage, alcohol and drug abuse, and binge-eating habits.

New Technology and Understimulation/Boredom

Rapid developments in computer technology and its application to industrial and commercial settings have been a major cause of stress in the workplace. A constantly changing environment is a potent stressor because employees must continually strive to keep up with new systems and methods. Technological change can bring stress in the form of social isolation, but one in which individuals are monitored. Jobs are much simplified because the computer-led process has taken over much of the hard physical effort and the decision-making process. Indeed Smith suggested that the main impact of visual display units on workplace activities was underload, de-skilling, and a lack of control. Nevertheless, the use of computers in the workplace has also resulted in responsibility being delegated further down the organizational chain than before, often without commensurate power or pay, thus causing other strains and pressures.

Conclusion

A job is boring if that job incumbent perceives it to be boring. Nevertheless, it is possible to design jobs to avoid such problems. For example, Persson et al. described this process as the design of active versus passive operator jobs, even when monitoring tasks are a predominant part of the job. To avoid problems associated with absenteeism or with employees leaving the organization in order to find a more fulfilling or stimulating job, it is possible to apply a

wide variety of human relations processes such as job enlargement, enrichment, restructuring, and rotation. As Zavala et al. suggested, the benefits are a reduction in boredom, work stress, absenteeism, and turnover and an increase in innovation, productivity, and loyalty. As such, these strategies are recommended as part of an organizational approach to stress management in the work environment.

See Also the Following Articles

Burnout; Workplace Stress.

Further Reading

Cheliout, F. (1979). Rythme thêta postêrieur au cour de la vielle active chez l'homme. *Rev EEC Neuropsychologie* **9**, 52–57.

Cooper, C. L. and Kelly, M. (1984). Stress among crane operators. *Journal of Occupational Medicine* **26**, 575–578.

Cox, T. (1985). Repetitive work: occupational stress and health. In: Cooper, C. L. & Smiths, M. J. (eds.) *Job stress and blue-collar work* pp. 85–112. Chichester, UK: Wiley.

French, J. R. P. and Caplan, R. D. (1973). Organizational stress and individual strain. In: Marrow, A. J. (ed.) *The failure of success*, pp. 30–66. New York: Amacon.

Granjean, E. (1968). Fatigue: its physiological and psychological significance. *Ergonomics* **11**, 427–436.

Greenglass, E. R. (2005). Proactive coping, resources and burnout: implications for occupational stress. In: Antoniou, A. S. G. & Cooper, C. L. (eds.) *Research companion to organizational health psychology* pp. 503–515. Cheltenham, UK: Edward Elgar.

ILO (1986). Psychosocial factors at work: recognition and control. *Report of the Joint International Labour Office and World Health Organization*, Ninth Session, 1984. Geneva: ILO.

Langan-Fox, J. (2005). New technology, the global economy and organizational environments: effects on employee stress, health and well-being. In: Antoniou, A. S. G. & Cooper, C. L. (eds.) *Research companion to organizational health psychology.* Cheltenham, UK: Edward Elgar.

Kass, S. J., Vodanovich, S. J. and Callender, A. (2001). State-trait boredom: relationship to absenteeism, tenure and job satisfaction. *Journal of Business and Psychology* **16**(2), 317–327.

Malinski, R. M. (2002). Job rotation in an academic library: dammed if you do and dammed if you don't. *Library Trends* **50**(4), 1–4.

Persson, A., Wanek, B. and Johansson, A. (2001). Passive versus active operator jobs in automated process control – a job design case study in a control centre. *Applied Ergonomics* **32**(5), 441–451.

Quick, J. C. and Quick, J. D. (1984). *Organizational stress and preventive management*. New York: McGraw-Hill.

Smith, M. J. (1997). Psychosocial aspects of working with visual display terminals (VDTs) and employee physical and mental heath. *Ergonomics* **40**(10), 1002–1015.

Sutherland, V. J. (2005). An organizational approach to stress management. In: Antoniou, A. S. G. & Cooper, C. L. (eds.) *Research companion to organizational health psychology.* Cheltenham, UK: Edward Elgar.

Sutherland, V. J. and Cooper, C. L. (1991). *Stress and accidents in the offshore oil and gas industry.* Houston, TX: Gulf.

Sutherland, V. J. and Cooper, C. L. (2000). *Strategic stress management: an organizational approach*. UK: Macmillan Business.

Sutherland, V. J. and Cooper, C. L. (2003). *De-stressing doctors: a self management guide*. London: Butterworth Heinemann.

Yerkes, R. M. and Dodson, J. D. (1908). The relation to the strength of the stimulus to the rapidity of habit formation. *Journal of Comparative Neurology and Psychology* **18**, 459–482.

Zavala, S. K., French, M. T., Zarkin, G. A. and Omachonu, V. K. (2002). Decision latitude and workload demand: implications for full and partial absenteeism. *Journal of Public Health Policy* **23**, 344–361.

Unemployment, Stress, and Health

M Bartley
University College London Medical School, London, UK

© 2007 Elsevier Inc. All rights reserved.

Introduction
Health Effects of Unemployment
Selection
Mechanisms
Unemployment and Health in the Life Course
Policy Implications

Glossary

Cohort study	A cohort study is one in which a group of individuals is followed over time. A birth cohort consists of a group whose members were all born at around the same time and are therefore the same age. The advantage of this for research on unemployment and health is that it allows us to see whether those who become unemployed differ in any health-relevant manner from those who do not, even before they experience unemployment.
Economically active	The employed and self-employed plus all those looking actively for paid employment.
Economically inactive	Persons who neither have a job nor are currently actively seeking one.
Employment	Work that either is paid by an employer or yields income from self-employment or a small business.
Latent consequences of employment	The idea that employment has benefits other than providing a wage, such as the opportunity to use skills, social integration, and self-esteem.
Life course perspective	A relatively new approach to understanding health differences between people in different social groups that takes account of experiences over time. For example, the approach looks at whether people who experience more unemployment tend to have had different experiences in childhood than those who have experienced little or no unemployment.
Mac-job	Employment that requires little education or training, is very low paid, and is accompanied by no career prospects, low autonomy, and low job security.
Parasuicide	This term does not mean attempted suicide that did not succeed, but refers to serious self-harm that is, however, not thought to be intended to result in death.
Reservation wage	The rate of pay at which an unemployed person will consider taking on a new job.
Selection	The possibility that pre-existing health may have resulted in unemployment.
Unemployment	Defined by the International Labour Organisation as "neither employed nor self employed and actively seeking employment." Practices in different nations vary as to the period of time during which an individual must have looked for a job in order to be regarded as unemployed; a common usage is within the last 2 weeks.
Unemployment rate	Calculated as the number of those seeking employment divided by the economically active population. Those not seeking paid work are not included in the denominator.
Work	Not synonymous with employment. It has been calculated that at least half of the work that is done in a society is unpaid. Most unpaid work is carried out by women.

Introduction

From levels in 1977 of around 4.5% in the UK and 6.5% in the US for men over the age of 16 in the mid-1970s, the unemployment rates of these nations rose sharply to around 11–12% in the early 1980s. In the UK, unemployment rose again to another peak, over 12% in men, in the early 1990s, while in the United States it fell more steadily. **Figures 1** and **2** show both similarities and differences in the trends in these two nations. Although unemployment is now far lower than its highest level, conditions are very different from those of the 1970s, which preceded the 1980s recession and restructuring of labor markets. The labor market has to some extent polarized. On the one hand, there are more jobs that require high levels of education and skill in new high-technology industries and, even more, in services such as banking, law, and health care. On the other hand, the jobs available to those without further post-school education and training have increasingly become what are known as Mac-jobs (after the McDonalds catering empire). The paradigmatic Mac-job is in the fast food industry and consists of relatively unskilled forms of preparing and serving of meals. Another job typical of the new labor market is the call center. Working in a call center does not require specific qualifications or skills, but is only possible for those capable of clear, grammatical speech, without a heavy regional dialect, and who have sufficient social skills to deal with many different

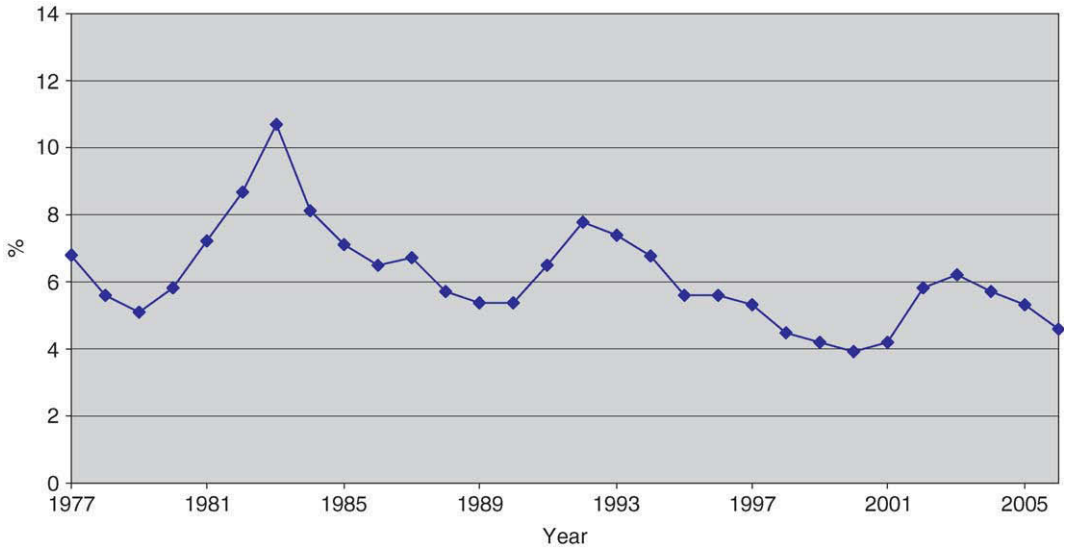


Figure 1 Trends in unemployment of working-age men (16 years and over), United States, 1977–2006 (January). From http://data.bls.gov.

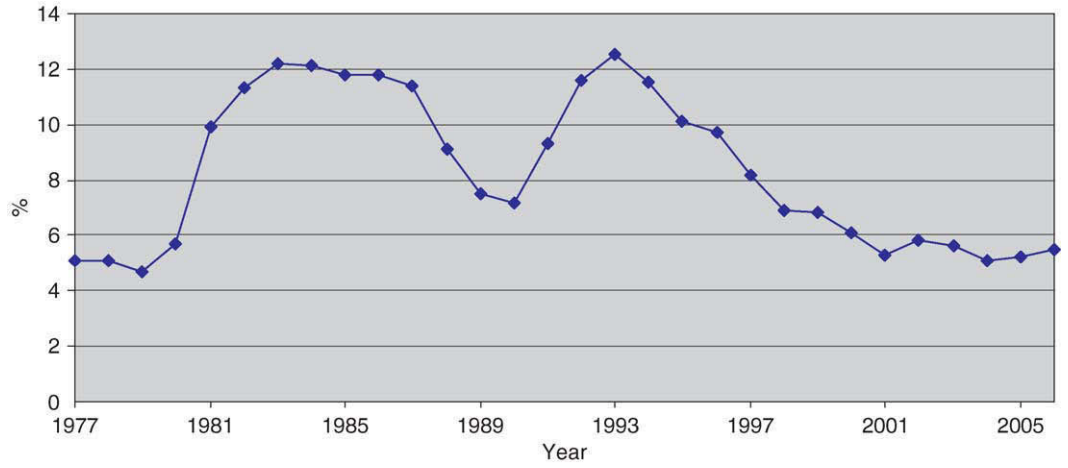


Figure 2 Trends in unemployment of working-age men (16 years and over), UK, 1977–2006 (January). From http://www.statistics.com.

(often dissatisfied) customers in a series of intensive remote encounters.

In many mainland European nations, however, large-scale replacement of unemployment by low-paid work has not yet taken place. Unemployment rates in Germany and France, for example, are still very high, at 10–11%. To some extent it could be said that mainland Europe and Scandinavia lag behind the United States and UK, in that the economic crisis of the 1980s in the latter nations occurred in mainland Europe and Scandinavia about 10 years later. It is, therefore, still relevant to consider whether there may be effects of unemployment on health.

Health Effects of Unemployment

Research has repeatedly shown a higher prevalence of ill health and excess mortality in men and women who are unemployed. Men who are unemployed, and women who are either unemployed or keeping house full time, are more likely than those with paid employment to describe their health as generally fair or poor. Damage to psychological health is also found by the vast majority of studies, an effect that appears to be independent of pre-existing health and to be reversed on re-employment in many cases.

However, it is not a simple step from this observation to a causal relationship between unemployment and health. In aggregate level studies, increases in unemployment in nations or areas are not consistently found to be related to increases in mortality or morbidity. The increase of unemployment in industrialized nations took place during a period when life expectancy was rising steadily. This is not too great a mystery, as even what would be regarded as a very high unemployment rate would be no more than

10–15% of the economically active population of a nation. A typical rate of economic activity in an industrial nation would be 65–75% of those regarded as of working age, which is usually between around 18 and 65 years old (the rest of the population comprising children, the retired, and those looking after home and family). It is also typical of observed patterns of unemployment that the risk of job loss is not randomly spread throughout the population. In general, there will be a group characterized by having worked in a declining industry, having few skills relevant to other industries, and living in an economically declining region, who experience long-term and/or multiple spells of unemployment. This group will carry most of the total burden of unemployment for the society as a whole. This needs to be remembered when we examine the evidence on the health effects of unemployment. It is more or less impossible to find a group of workers who become and remain unemployed for any appreciable period of time who do not also suffer from one or more other burdens of socioeconomic disadvantage.

The reason we need to be cautious about studies showing that people who report that they are unemployed also report worse mental or physical health is that those who are ill may be more likely to lose their jobs and find it harder to regain employment Or there may be indirect effects whereby men and women with other characteristics such as lower levels of education, which are known to be related to health risks, are also less likely to be employed. This type of relationship, in which certain people are more at risk of both unemployment and ill health is known as selection.

Selection

There are very few studies that have examined relationships of unemployment to health over a long time period. Convincing evidence of a causal effect would be present if long-term studies showed that healthy people who became unemployed subsequently experienced a deterioration in their health. Early studies of unemployed people in the 1970s and 1980s indicated that their physical health did not necessarily decline during a spell of unemployment; a number of studies failed to find an increase in morbidity among those who were continuously unemployed for up to 18 months. However, a more recent study of a representative sample of healthy people over a period of 10 years did observe an increased risk of developing a limiting illness in the year after having become unemployed. Another careful study carried out by occupational health specialists measured a clear deterioration in physical working capacity after a year of unemployment.

However, because of the scarcity of good long-term evidence on the health of people who become unemployed, we need to consider the possibility of what is called a selection effect. A direct health selection effect is thought to be at work when it is poorer health itself that increases the risk of unemployment. Ill health may be a risk for both initial job loss and subsequent chances of re-employment. However, studies in which an entire workplace closes down and all workers are dismissed so that there is no possibility for job loss to be affected by health have also shown increases in illness in those made redundant. Research using linked census-based data from England and Wales and from the Nordic countries has shown that higher risk of mortality is not only present among unemployed people who were already ill. A recent study in Sweden has also found excess risk of mortality among twin siblings who experienced unemployment when compared to their siblings who did not.

Although most studies of increases in national or regional unemployment do not find that these are followed by deteriorations in population health, there is most certainly no evidence that increases in unemployment happen because people's health gets worse. The large increases in unemployment during the 1980s in the Anglo-Saxon nations and during the 1990s in many European nations were not due to worsening health of workers in these countries (indeed, this will seem rather an absurd statement to make). In fact, unemployment rates increased dramatically in different nations during the 1980s and 1990s, at the same time that levels of health in the population actually improved. So there is no possibility that an increase in unemployment overall can be blamed on some kind of general deterioration in the health of the working-age population.

There may still be some tendency for those who remain unemployed for longer periods to be different from those who do not, even if this difference does not take the form of a life-threatening disease. Perhaps people at high risk of unemployment may also have certain personality characteristics, for example, an external locus of control or a weak coping ability. However, this is a more complex hypothesis, and is not necessarily the same as selection. These ideas have to be considered alongside other evidence on the life course accumulation of social and health disadvantage.

One study that has been able to ask the question "How far is the poor health of the unemployed due to their health before they became unemployed?" in relation to mental health is a British birth cohort study of everyone born in the second week of March 1958 and followed until the present time. Researchers

were able to use data on health and psychological development throughout the school years and young adult life, as well as complete employment histories from leaving school to age 33. These data showed that cohort members who experienced more unemployment were also more likely to have grown up in economically disadvantaged and overcrowded households. The study also addressed the possibility that the high level of psychological morbidity found in unemployed men and women might have been due to a pre-existing vulnerability to poor mental health. Even after taking into account pre-existing mental health, recent unemployment was clearly related to the onset of psychiatric symptoms severe enough to require medical care. However, when those with a prior tendency to depression were excluded from the analysis, this had the effect of strengthening the relationship between longer term accumulation of unemployment and the onset of episodes of depression and anxiety. This adds to the evidence that longer term unemployment causes deterioration in mental health in those who were previously healthy.

Mechanisms

What might be the reasons for the relationship between unemployment and health? Several have been put forward, and the following sections consider three of these: poverty, the fact that unemployment is a stressful life event, and changes in health-related behaviors at the time of unemployment.

Poverty

Low living standards are not an inevitable consequence of unemployment; this is a result of the levels at which benefits are set, which results from political decisions. During the 1980s and 1990s, levels of income replacement for the unemployed were lowered in many nations. It was argued that under conditions of increasing automation of unskilled and semi-skilled work, levels of benefits available to the unemployed exceeded the market worth of their labor, that is, the wage rate that employers were willing to pay. If the level of pay at which an unemployed person will accept new employment – the reservation wage – is too high, employers will not take them on. If benefit levels are too high, the state is raising the reservation wage and therefore may be contributing to the problem of high unemployment.

On becoming unemployed, most people will experience a decrease in income. The extent of this decrease varies between nations, from around 20% in some Nordic countries to approximately 60–70% or more in the UK and United States. Many studies link the health effects of unemployment directly to financial problems. In modern welfare states, state benefits available to those unable to undertake paid employment are thought to avoid starvation and severe physical privation. However, recent research on the financial cost of what medical evidence has shown to be a healthy lifestyle, which includes a diet adequate in essential nutrients, warm housing, and a level of social participation consistent with mental health, costs approximately double the level of social security benefit in the UK. It is therefore inevitable that once savings are exhausted, unemployed people will be forced to borrow money in order to sustain health. Debt, in its turn, has been shown to increase the risk to mental health and the risk of stress-related deterioration in physical health. In several studies, long-term unemployed people who had gone into debt or had to borrow money in the past year had a risk of depression that was more than double that of those who did not have to borrow money. These studies also found that those obliged to borrow were also more likely to report deterioration in physical health. Others have documented the ways in which increasing financial pressures, as savings are used up and worn-out items need to be replaced, are responsible for the growing inactivity and social isolation of many unemployed people.

These British findings are echoed in other countries, such as the United States and the Netherlands, where financial problems were shown to be the main reason why unemployed people's mental and physical health were worse than that of the employed. However, it seems that eventually many of the unemployed adapt to straitened financial and social circumstances. Several studies agree that there appears to be no further deterioration in psychological well-being after a period of between 1 year and 18 months (but no improvement, either, so that unemployed people remain two to three times more likely than the employed to be in poor mental health). This adds weight to the argument for providing early assistance to those who become unemployed. Other evidence suggests that adaptation to unemployment is accompanied by lowered expectations of oneself, and perhaps by a degree of alienation and cynicism. This process has been compared to the adaptive process found in institutionalized inmates of prisons or psychiatric hospitals.

Unemployment as a Stressful Life Event

Research on the ways in which people react to job loss shows it to be a highly stressful life event, which has been characterized as a form of bereavement. International studies provide consistent evidence of the importance of stress. Many researchers have suggested that employment has a number of nonfinancial

benefits to the individual, the so-called latent consequences, and it is the loss of these that results in the threatening character of unemployment. These latent consequences of employment include giving a time structure to the day, self-esteem, and the respect of others; physical and mental activity; use of skills; and interpersonal contact. It may be that the loss of these psychosocial benefits of employment is at least as important as loss of income to the health of unemployed people.

Evidence of the importance of stress for the relationship between unemployment and health comes from studies in Scandinavian nations. In these nations, cash benefits to the unemployed are relatively generous, and therefore the financial effects might not be expected to be as great as in the United States or the UK; however, similar relationships between unemployment and a range of health indicators are seen. A study in Stockholm of men with irregular work histories and a variety of problems that required frequent social service assistance found that those who were employed, even in low-paid casual jobs, were all more active and integrated (and psychologically healthier) than the unemployed. Among unemployed industrial workers in Finland, those who regained paid work experienced a considerable improvement in psychological health regardless of their financial circumstances either before or after re-employment. Italian workers laid off from their jobs experienced raised amounts of both psychological and physical illness despite receiving the whole of their normal wage. These studies provide evidence of the nonfinancial benefits of employment for psychological health. Stress may affect physical health as well as psychological health as a result perhaps of chronically increased levels of anxiety and the relationship this may have to the functioning of the endocrine system.

Health-Related Behavior

There is evidence that unemployment is associated with some forms of health-damaging behavior, although previous research has yielded inconsistent findings on the relationship. Associations between unemployment and health behaviors need to be seen in the context of the longer term development of these behaviors. The little research that exists on this topic seems to indicate not so much that people who become unemployed take up new, health-damaging activities such as smoking and drinking, not least because the costs of these pursuits could be ill-afforded. Rather, it seems that the experience of unemployment makes it harder for individuals to carry through their intentions to improve their health behaviors. For example, in one study, after taking account of adolescent health risk behaviors, even well-educated and intelligent young men who experienced 12 months of more without a job between the ages of 23 and 33 were less likely than their peers with no unemployment experience to have a healthy lifestyle in their mid-30s.

Self-destructive behavior among unemployed men has been widely investigated. In the UK and in Scandinavian nations, suicide has been found to be more likely in unemployed people. Parasuicide (attempted suicide) has also been found to be higher in unemployed men. Unemployment seems to increase the likelihood of other adverse life events and to reduce the psychological and social resources needed to cope with these. There is, for example, a higher risk of partnership and marriage breakdown among unemployed men. Many of the unemployed also lose contact with their social networks due to the stigma of their situation.

Unemployment and Health in the Life Course

Perhaps the best way to approach the study of the relationship between health and labor market status is from the perspective that unemployment and work insecurity are part of a process through which health disadvantage is accumulated over the life course. Research taking this perspective turns away from any simple opposition between selection and causation and shows that unemployment occurs as part of a much longer term sequence of events in the life of the individual. It is now becoming possible to look at these processes using cohort studies that follow individuals from childhood to middle age. We can see, for example, that those most likely to experience unemployment may also be more vulnerable to excess mortality and morbidity because of earlier experiences of other kinds of hardship and disadvantage. Their experience of unemployment itself therefore contributes to a process of accumulation of health risks. There is little research conducted along these lines. One such analysis showed that while those with disadvantaged childhoods had the highest risk of later unemployment during the recession of the 1980s, young men who had experienced over 12 months of unemployment seemed to be on a less healthy track in terms of both social and behavioral risk, even if they were from more advantaged social origins. It seems as though unemployment also tends to wipe out some of the health advantages built up over the period from childhood to young adulthood in those who had come from socially advantaged homes and done relatively well in education.

Recent longitudinal research has made it possible to understand the relationship between unemployment and health in a more dynamic way. Biological endowment and material and emotional experiences during childhood confer varying amounts of health resources or health potential to the young person. In the real world, people from less secure and advantaged backgrounds are more likely to enter occupations with more adverse working conditions and lower rates of pay. It is of no surprise, therefore, that the aging process itself is known to take place at differential rates according to the employment history of the individual, with more arduous work being associated with faster aging. The physical resources the individual brings to the labor market when he or she must look for a new job will in many cases have already been damaged by these experiences at work, even before any spell of unemployment he or she may suffer. Many of the most hazardous (and least well paid) occupations also tend to be less secure, which means that there will be some degree of coincidence between poor health and the risk of job loss, not just because of the individual's own characteristics but also because of social structure. In contrast to physical resources, psychological and social resources are more likely to be enhanced by the experiences of activity, companionship, and cooperation involved in (even hazardous) employment. Deterioration in psychological health during unemployment is therefore more immediate and visible than deteriorations in physical health, and improvement on re-employment happens quickly.

For young people, employment forms part of the process of establishing an independent identity, so that stable employment may in fact be even more important for younger than for middle-aged or older people. For many, entry into stable employment is now preceded by long periods of job insecurity. The process of identity formation is therefore far more at risk, with accompanying rises in rates of relationship breakdown, poor mental health, addiction, and accidental and self-inflicted harm. Employment is known to aid the development of secure identity and self-esteem and to facilitate the formation of stable relationships. Accordingly, as the British economy shifted from one in which there were larger numbers of stable jobs to one in which employment conditions were more insecure, British birth cohort studies showed that levels of mental health problems in those of young working age increased sharply. Although rates of major diseases of middle age have continued to decline during the period of labor market change, rates of suicide in young men have continued to rise.

Policy Implications

A growing body of international comparative research suggests that societies with higher levels of unemployment and job insecurity may actually produce different sorts of life histories in individual members than those that provide larger numbers of secure jobs. It has been all too easy to take for granted some of the effects of quasi-full employment on the wider society. When jobs were plentiful, in order to recruit and retain workers most large firms undertook extensive training, recruiting school leavers at an early age into apprenticeships. Schools therefore had far fewer restless and unmotivated youth to deal with than would otherwise have been the case, benefiting all children in the school. Internal labor markets (the availability of promotion within a firm based on experience rather than on qualifications) offered people the chance to advance in their jobs even in the absence of school qualifications, from a later starting point. Increasing numbers of firms offered occupational pensions, thus raising the living standards of the retired population at little cost to the state. As part of the McDonaldization of the labor market, the costs of training and of supporting those who are no longer able to work for payment are now increasingly being borne by the individual (the disappearance of apprenticeships and the appearance of student loans, critical illness policies, and private pensions are examples). There has also been an increase in the numbers of persons who become economically inactive, which in many countries far exceeds the increase in the numbers of those officially unemployed and seeking employment.

Increasing the number of low-paid, low-skilled Mac-jobs is often advocated as a solution to the problem of unemployment. However, the evidence on the success of such a policy is mixed, and the impact on population health has not as yet been fully investigated.

During the 1990s, new data from birth cohort studies (studies of large numbers of people all born in the same year) added a considerable amount to our understanding of why unemployment is associated with poor health. We began to see that unemployment takes place as part of a process of accumulation of disadvantage that may begin in childhood, and that a spell of unemployment often occurs as part of a more general pattern of hazardous and insecure work. The old assumptions behind the notion of a social insurance or welfare state safety net are seen to be too simple. The people most likely to fall into the net bear a heavier weight of disadvantage than those who are less likely to need its support. The idea that benefit levels for the unemployed should be set no

higher than mere subsistence, in order not to encourage voluntary unemployment, is therefore becoming discredited. On the contrary, both the taxpayer and the unemployed person (who is of course a past and potential taxpayer him- or herself) benefits more from a generous provision of both material and psychosocial support for those who find themselves without employment.

See Also the Following Articles

Health Behavior and Stress; Immune Surveillance – Cancer, Effects of Stress on; Job Insecurity: The Health Effects of a Psychosocial Work Stressor.

Further Reading

Allmendinger, J. and Hinz, T. (1998). Occupational careers under different welfare regimes: West Germany, Britain and Sweden. In: Leisering, L. & Walker, R. (eds.) *The dynamics of modern society*, pp. 64–84. Bristol, UK: Policy Press.

Bethune, A. (1996). Economic activity and mortality of the 1981 Census cohort in the OPCS Longitudinal Study. *Population Trends* **83**, 37–42.

Burchell, B. (1996). Who is affected by unemployment? In: Gallie, D., Marsh, C. & Vogler, C. (eds.) *Unemployment and social change*, pp. 45–58. Oxford, UK: Oxford University Press.

Catalano, R., Dooley, D., Wilson, G. and Hough, R. (1993). Job loss and alcohol abuse: a test using data from the Epidemiologic Catchment Area project. *Journal of Health Society and Behavior* **34**, 215–225.

Catalano, R., Dooley, D., Novaco, R. W., Wilson, G. and Hough, R. (1993). Using ECA survey data to examine the effect of job layoffs on violent behavior. *Hospital and Community Psychiatry* **44**, 874–879.

Dooley, D. and Catalano, R. (1991). Unemployment as a stressor: findings and implications of a recent study. *WHO Regional Publications European Series* **37**, 313–339.

Dooley, D., Fielding, J. and Levi, L. (1996). Health and unemployment. *Annual Reviews of Public Health* **17**, 449–465.

Kessler, R. C., Turner, J. B. and House, J. S. (1987). Intervening processes in the relationship between unemployment and health. *Psychological Medicine* **17**, 949–961.

Levenstein, S., Smith, M. W. and Kaplan, G. A. (2001). Psychosocial predictors of hypertension in men and women. *Archives of Internal Medicine* **161**, 1341–1346.

Montgomery, S. M., Bartley, M. J., Cook, D. G. and Wadsworth, M. E. (1996). Health and social precursors of unemployment in young men in Great Britain. *Journal of Epidemiology and Community Health* **50**, 415–422.

Montgomery, S. M., Cook, D. G., Bartley, M. J. and Wadsworth, M. E. J. (1998). Unemployment, cigarette smoking, alcohol consumption and body weight in young British men. *European Journal of Public Health* **8**, 21–27.

Montgomery, S. M., Cook, D. G., Bartley, M. J. and Wadsworth, M. E. (1999). Unemployment pre-dates symptoms of depression and anxiety resulting in medical consultation in young men. *International Journal of Epidemiology* **28**, 95–100.

Morris, J. N., Donkin, A. J. M., Wonderling, D., Wilkinson, P. and Dowler, E. A. (2000). A minimum income for healthy living. *Journal of Epidemiology and Community Health* **54**, 885–889.

Moser, K., Goldblatt, P. O., Fox, A. J. and Jones, D. R. (1987). Unemployment and mortality. *British Medical Journal* **294**, 509–512.

Moser, K., Goldblatt, P., Fox, J. and Jones, D. (1990). Unemployment and mortality. In: Goldblatt, P. O. (ed.) *Longitudinal study: mortality and social organisation*, pp. 77–102. London: HMSO.

Ostry, A. S. (2001). Effects of de-industrialization on unemployment, re-employment, and work conditions in a manufacturing workforce. *BMC Public Health* **1**, 15.

Ritzer, G. (1993). *The McDonaldization of society*. Oakland, CA: Pine Forge Press.

Wadsworth, M. E., Montgomery, S. M. and Bartley, M. J. (1999). The persisting effect of unemployment on health and social well-being in men early in working life. *Social Science and Medicine* **48**, 1491–1499.

Warr, P. (1987). *Work, unemployment and mental health*. London: Oxford University Press.

Urocortins

É M Fekete and E P Zorrilla
The Scripps Research Institute, La Jolla, CA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E P Zorrilla and G F Koob, volume 3, pp 637–642, © 2000, Elsevier Inc.

Structure of Urocortins, CRF, and Related Peptides
Distribution of Urocortins
Pharmacology of Urocortins
Physiological and Behavioral Effects of Urocortins

Glossary

Adrenocorticotropic hormone (ACTH)	A hormone released into systemic circulation by the anterior pituitary as part of the HPA-stress response. ACTH, or corticotropin, induces the release of glucocorticoids from the adrenal cortex.
Corticotropin-releasing factor (CRF)	Isolated by Vale and colleagues in 1981, the first hypothalamic ACTH secretagogue to be identified. In addition to its hypothalamic endocrine function, extrahypothalamic CRF is important in behavioral and autonomic stress responses.
Edinger-Westphal nucleus	An oculomotor nucleus of the midbrain that has parasympathetic projections to the ciliary ganglion and descending inputs to the brain stem and spinal cord.
Hypothalamic-pituitary-adrenocortical (HPA) axis	The neuroendocrine system which, during stress, amplifies a neural signal of physiological or psychological stress into a systemic, endocrine response. Hormones secreted as part of the HPA stress response cascade equip the organism to mobilize and utilize energy resources more effectively and, more generally, respond to the stressor.
Stresscopin	An N-terminally extended analog of urocortin 3 that binds exclusively and with high affinity to the CRF_2 receptor.
Stresscopin-related peptide	An N-terminally extended analog of urocortin 2 that binds selectively and with high affinity to the CRF_2 receptor.

Since the isolation of corticotropin-releasing factor (CRF) in 1981, three more endogenous CRF-like peptides have been identified in mammals. The first of these – urocortin 1 (Ucn 1) – was identified in 1995 for its similar primary structure and bioactivity to suckerfish urotensin I and CRF. Ucn 1 exhibited greater affinity for and activation of type 2 CRF receptors (CRF_2) than CRF and was hypothesized to be a natural CRF_2 receptor ligand. Urocortin 2 (Ucn 2) and urocortin 3 (Ucn 3), 38-amino-acid residue peptides, were then identified in 2001 for their structural relation to both CRF and Ucn 1 and were named urocortins due to their predominant affinity for the CRF_2 receptor. Although the urocortins share sequence identity with one another and with CRF, each is phylogenetically distinct, with the CRF peptide family resulting from multiple gene duplication events. Each peptide has a unique anatomical distribution under the control of different genes. Consequently, despite a structural family resemblance, the natural functions of urocortins and CRF in stress responses *in vivo* may differ significantly.

Structure of Urocortins, CRF, and Related Peptides

The rat Ucn 1 cDNA was cloned from a rat midbrain mRNA library in 1995. The gene codes for a 122-residue prohormone that contains Ucn 1 in the C-terminus. Ucn 1 genes also have been cloned in mouse, human, hamster, sheep (*Ovis aries*), frog (*Xenopus laevis*), and capuchin monkey (*Cebus apella*) and are similar across species. The human Ucn 1 gene lies on chromosome 2 (2p23-p21) and has two exons, with the second exon comprising the coding region, like the CRF gene. The Ucn 1 promoter is regulated by several transcription factor binding sites, including a TATA box, GATA-binding sites, a C/EBP-binding site, a Brn-2 binding site, and a cyclic AMP responsive element (CRE). Upstream of the CRE site (four base pairs), the Ucn 1 promoter contains a consensus half site for glucocorticoid response elements (GREs), consistent with the finding that glucocorticoids upregulate Ucn 1 mRNA synthesis.

The cDNAs for Ucn 2 and Ucn 3 were identified from sequence homology searches of the mouse and human genomes, with the putative Ucn 2 peptide (an N-terminally shortened sequence of stresscopin-related peptide) contained in the C-terminus of a 112-amino-acid residue proprotein, and the putative Ucn 3 peptide (an N-terminally shortened sequence of stresscopin) encoded in a 161-residue precursor. The identity of the endogenous peptides derived from the Ucn 2 and Ucn 3 preproteins remains predicted, as mature peptide sequences have not been definitively isolated and sequenced. Whether a mature Ucn 2 peptide sequence is processed in humans remains uncertain because of the lack of a consensus proteolytic

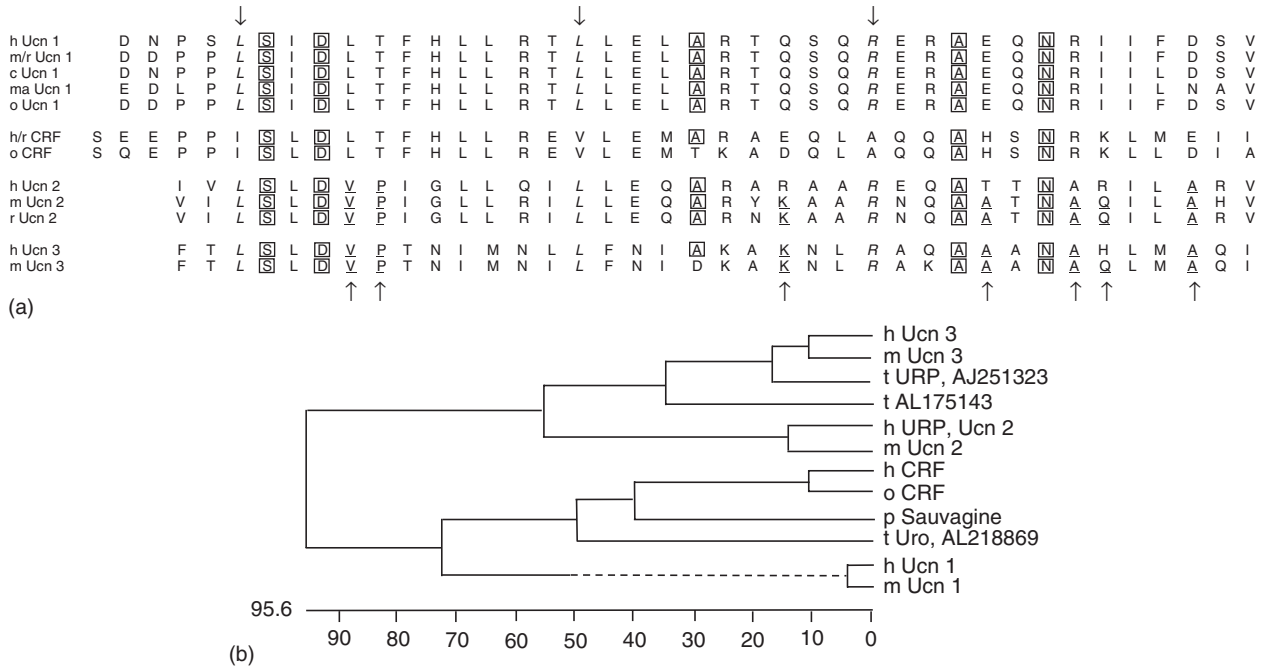


Figure 1 (a) Comparison of the primary structures of mammalian CRF family peptides. Selected putative amino acid sequences for CRF and urocortins 1, 2, and 3 across human (h), murine (*Mus musculus* [m]), rat (*Rattus norvegicus* [r]), monkey (*Cebus paella* [c]), hamster (*Mesocricetus auratus* [ma]), and sheep (*Ovis aries* [o]). Boxed regions indicate CRF superfamily homology; residues in italics and marked by down arrows are shared among all the urocortins; underlined residues marked by up arrows are shared among the type 2 urocortins (Ucn 2 and Ucn 3), but not by Ucn 1. (b) The phylogenetic tree associates human and mouse Ucn 3 and Ucn 2 with pufferfish (*Takifugu rubripes* [t]) urocortin-related peptide (URP). The more distantly related lineage is composed of ovine and human CRF, human and mouse Ucn 1, frog (*Phyllomedusa sauvagei* [p]) sauvagine, and pufferfish urotensin (tUro). The scale beneath the tree reflects sequence distances. Genbank accession sequences are provided for pufferfish-derived peptides. The phylogenetic tree was adapted from Lewis et al. (2001). *Proceedings of the National Academy of Sciences USA* **98,** 7570–7575.

processing site at the C-terminus of the putative sequence. Preproteins for rat Ucn 2 and for chimpanzee and frog Ucn 3 have also been identified. The human genes for Ucn 2 and Ucn 3 lie on chromosomes 3 (3p21.3) and 10 (10p15.1), respectively. The Ucn 2 promoter contains several GREs and is positively regulated by glucocorticoids.

Ucn 1 resembles CRF in primary structure (see Figure 1) more so than it resembles Ucn 2 and Ucn 3. In contrast, Ucn 2 and Ucn 3 resemble one another more than they do CRF, a distinction, coupled with their greater CRF_2 receptor selectivity, that has led to their characterization as type 2 urocortins. An ancient gene duplication event is believed to have given rise to separate Ucn 1/CRF vs. Ucn 2/Ucn 3 lineages, with subsequent gene duplications events differentiating paralogs within each evolutionary branch (see Figure 1).

Distribution of Urocortins

Central Nervous System

Ucn 1 has a restricted, subcortical, predominantly caudal distribution in brain. The major site of brain Ucn 1 synthesis is the Edinger-Westphal nucleus (E-W), a dorsal midbrain structure known for its role in oculomotor function. The prominent synthesis of Ucn 1 in the E-W is conserved across rats, sheep, humans, and frogs. Descending Ucn 1-like immunoreactive (LI) positive fibers of possible E-W origin are observed in (1) midbrain: substantia nigra, periaqueductal gray, interpenduncular nucleus, and red nucleus; (2) caudal midbrain/rostral pons: dorsal raphe nucleus, ventral tegmentum, basilar pontine nuclei, and parabrachial nucleus; (3) medulla, including the facial, lateral reticular, and spinal trigeminal nuclei, inferior olive, and the dorsal column nuclei, nucleus of the solitary tract (NTS), and area postrema; (4) cerebellum: in the flocculus and paraflocculus as well as deep cerebellar and vestibular nuclei; and (5) spinal cord: throughout the spinal gray and, less so, in the dorsal and ventral horns. Ascending Ucn 1-LI-positive projections, possibly from the E-W, target the septal/preoptic region and, less heavily, the hypothalamus and thalamus. Validated secondary sites of brain Ucn 1 synthesis include the lateral superior olive, the supraoptic nucleus of the hypothalamus (SON), the lateral hypothalamic area, and several brain stem and spinal cord motoneuron nuclei. Ucn 1-LI is scarce in

many regions in which CRF is prominent, including the external layer of the median eminence, the hypophysiotropic, dorsal medial parvocellular subdivision of the paraventricular nucleus of the hypothalamus (PVN), basal ganglia, amygdala, hippocampus, locus coeruleus (LC), and cerebral cortex.

Ucn 2 also exhibits a restricted, subcortical expression in rodent brain. Like Ucn 1, Ucn 2 mRNA is present in the SON and magnocellular subdivision of the PVN, in brain stem motoneurons, and in the spinal cord. Unlike Ucn 1, Ucn 2 also is expressed in the arcuate nucleus of the hypothalamus and the LC. The projection targets of Ucn 2 neurons are unknown. Nonneuronal Ucn 2 expression also is present in meninges, but not glia.

Ucn 3 exhibits the most rostral, but again subcortical, distribution of the urocortins. The three main sites of forebrain Ucn 3 synthesis are (1) the median preoptic nucleus of the hypothalamus, (2) a perifornical hypothalamic region bordered laterally by the fornix and medially by the PVN, and (3) the dorsal medial amygdala. Less prominent forebrain sites of Ucn 3 synthesis include the dorsomedial hypothalamus, both magnocellular and parvocellular components of the PVN, a region dorsal to the SON, and the posterior cortical and amygdalohippocampal transition areas of the amygdala. Ucn 3-LI-positive fibers of uncertain origin project heavily to the ventromedial nucleus of the hypothamalus (VMH) and arcuate nucleus and less so to the ventral premammillary nucleus. Outside the hypothalamus, forebrain Ucn 3 fibers of uncertain origin are abundant in the lateral septum, posterior bed nucleus of stria terminals (BNST), and the medial amygdala. The topographical distribution of Ucn 3 in the LS differs from that of Ucn 1, with the latter innervating ventral and intermediate aspects of the structure, and the former innervating dorsal aspects. Caudally, Ucn 3 cell bodies are found in the auditory complex, and scattered Ucn 3 fibers are present in the periaqueductal gray, superior and inferior colliculi, and the ventral lateral lemniscus.

Periphery

Urocortins also are distributed in the periphery. Ucn 1 expression has been observed in adipose tissue, heart, and immune tissue, including thymus, spleen, and skin. Ucn 1 also is present in the enteric nervous system of the duodenum, small intestine, and colon, as well as in testis, kidney, adrenals, pancreas, and, possibly, anterior (but not posterior) pituitary. Ucn-1-LI also has consistently been observed in parietal and oxyntic cells of the stomach. Whether this Ucn 1 is actually synthesized by gastric tissue remains uncertain, however, as Ucn 1 is synthesized by lamina propria macrophages, components of the stomach's inflammatory mucosal immune system. Finally, Ucn 1-LI is abundant in human placenta and fetal membranes.

In humans, Ucn 2 gene expression was detected in most peripheral tissues analyzed by polymerase chain reaction analysis, with the highest levels observed in heart, lung, muscle, stomach, and adrenal and peripheral blood cells, but not skin. In the mouse, Ucn 2 gene expression was high in skeletal muscle and skin (unlike humans), moderate in lung, stomach, adrenal, ovary, brown fat, spleen, and thymus, and low in uterus, testes, kidney, white fat, and intestine. Unlike in humans, only low Ucn 2 mRNA expression is seen in murine heart. Skin, but not skeletal muscle, expression of Ucn 2 mRNA in the mouse appears to be inversely related to circulating glucocorticoid levels.

Ucn 3 gene expression is present in adipose tissue, heart, muscle, adrenal gland, and skin, albeit at levels considerably lower than those of Ucn 2. Ucn 3 also is present in thyroid, adrenals, β cells of the pancreas, spleen, and the muscularis mucosa of the gastrointestinal (GI) tract, especially in the stomach, small intestine, and colon, but not esophagus or rectum.

Pharmacology of Urocortins

Ucn 1 binds with high affinity to every mammalian CRF binding site identified to date, including the CRF_1 and CRF_2 receptor families and the CRF-binding protein (CRF-BP). This nonselectivity differs from that of the type 2 urocortins, which have selective CRF_2 affinity and less affinity for the CRF-BP. Three CRF receptor subtypes have been identified (CRF_1, CRF_2, CRF_3), two of which have been observed in mammals (CRF_1, CRF_2). CRF receptors are encoded by separate genes, share high sequency homology, and have unique tissue distributions and pharmacological affinity profiles, permitting a diversity of function. CRF receptors belong to the G-protein-coupled, seven transmembrane domain receptor (GPCR) family.

CRF_1 Receptors

The CRF_1 receptor is a class B (secretin-like) GPCR spanning 415 amino acids. Many alternatively spliced transcripts of CRF_1 have been identified in humans, rats, mice, and hamsters. However, only one of these splice variants is known to induce signal transduction in humans and rodents ($CRF_{1(a)}$), so the following pharmacological characterization pertains mainly to the $CRF_{1(a)}$ isoform. Ucn 1 exhibits reversible, saturable, high-affinity binding to CRF_1 receptors

transfected in stable cell lines. Ucn 1 is two to six times more potent than CRF at binding to the CRF_1 receptor and two- to threefold more potent than CRF in stimulating the production of cAMP from CRF_1 expressing cells *in vitro*. Ucn 1's affinity for the CRF_1 receptor is determined by regions in the first three extracellular domains of the receptor. In contrast to Ucn 1, the type 2 urocortins have very low potencies to activate adenylate cyclase via CRF_1 (EC_{50} 100 nM), and Ucn 3 shows negligible potency to stimulate ACTH release *in vitro* (1 μM), a CRF_1-mediated bioassay. The CRF_1 receptor can employ multiple signal transduction pathways when stimulated by Ucn 1. These include activation of adenylate cyclase and protein kinase A-dependent pathways, activation of phospholipase C and protein kinase C-dependent pathways, mitogen-activated protein (MAP) kinase-dependent pathways, nitric oxide production, and proximate interactions with calcium channels.

CRF_2 Receptors

CRF_2 receptors also are class B GPCRs. To date, four major CRF_2 splice variants have been identified, including membrane-bound and soluble isoforms of $CRF_{2(a)}$ and membrane-bound $CRF_{2(b)}$ and $CRF_{2(c)}$ receptors. The membrane-bound CRF_2 receptors are, respectively, 411, 431, and 397 residues in length, differing only in their extracellular N-terminus, with the $CRF_{2(c)}$ receptor observed thus far only in human limbic circuits. Unlike CRF, all urocortins bind with high affinity to membrane-bound CRF_2 receptors transfected in stable cell lines. Ucn 1 is slightly more potent than the type 2 urocortins, but is 10–40 times more potent than CRF at binding to membrane CRF_2 receptors. However, the type 2 urocortins are much more (1000-fold) selective CRF_2 agonists than Ucn 1. The CRF_2 receptor also can activate the MAP kinase pathway, with some evidence suggesting that Ucn 3 may be less efficacious than Ucn 2 and, especially, Ucn 1 at engaging this signal transduction mechanism. The soluble $CRF_{2(a)}$ isoform was identified in mice and results from a frameshift deletion of exon 6. In contrast to membrane-bound CRF_2 receptors, the soluble $CRF_{2(a)}$ receptor unexpectedly shows high affinity for Ucn 1 and CRF (classically, CRF_1 ligands) and low affinity for Ucn 2 and Ucn 3. Similar to the CRF-BP (see below), the soluble $CRF_{2(a)}$ receptor may curb CRF_1 signaling by competitively sequestering ligand.

CRF-Binding Protein (CRF-BP)

The CRF-BP is a conserved 37-kDa secreted glycoprotein that binds CRF and Ucn 1 with high affinity, thereby potentially reducing their bioavailability. Ucn 1 and CRF are approximately equipotent at binding to the CRF-BP. Ucn 1 can displace CRF from the CRF-BP (and vice versa), thereby allowing the possibility that each ligand modulates levels of free CRF receptor agonist in the brain. The region of Ucn 1 that has affinity for the CRF-BP (residues 4–28) differs from that which underlies its affinity for CRF receptors. Ucn 1 is the main natural ligand for the CRF-BP in sheep brain and dissociates from the CRF-BP more slowly than CRF. In contrast to CRF and Ucn 1, mUcn 3 does not bind with high affinity to the human or rat CRF-BP, and mUcn 2 binds with an order lower affinity than CRF or Ucn 1 to the rat CRF-BP and with negligible affinity to the human CRF-BP.

Physiological and Behavioral Effects of Urocortins

Hypothalamic-Pituitary-Adrenocortical (HPA) Axis

Although exogenous administration of Ucn 1 can potently activate the HPA axis, it is doubtful that endogenous urocortin 1 is a physiological regulator of the HPA axis. Unlike CRF-deficient mice, Ucn 1-deficient mice exhibit normal basal and stress-induced HPA hormone levels. Similarly, unlike CRF antisera, peripheral administration of specific Ucn 1 antisera does not modify basal, stress-induced, or adrenalectomy-induced ACTH levels. Finally, unlike the distribution of CRF, Ucn 1-immunoreactive fibers are scarce in the PVN and the external layer of the median eminence under basal conditions. Unlike both CRF and Ucn 1, intravenous Ucn 2 and Ucn 3 do not increase ACTH or corticosterone secretion, consistent with the absence of CRF_2 on pituitary corticotroph cells. However, type 2 urocortins might locally modulate CRF synthesis at the hypothalamic level, because immobilization stress increases Ucn 2 and Ucn 3 mRNA in the parvocellular PVN.

Osmoregulation

Because Ucn 1 and Ucn 2 are synthesized by magnocellular neurons of the SON and because Ucn 1 fibers are present in the posterior pituitary, an osmoregulatory role for urocortins is hypothesized. Supporting this possibility, salt loading, dehydration, and hypophysectomy increase Ucn-LI in magnocellular SON and PVN neurons, and food deprivation decreases Ucn-LI in the SON. Salt loading also increases SON Ucn 1 mRNA expression. Finally, chronic osmotic stimulation increases CRF_2 mRNA levels in the SON and magnocellular PVN, potentially increasing sensitivity to local urocortin action. Collectively, the

findings support a potential endogenous role for central urocortins in the regulation of salt/water balance. Peripheral urocortins also may control body fluid homeostasis as a result of their ability to stimulate atrial natriuretic peptide (ANP) release from cardiomyocytes, again through a putative CRF_2 mechanism. ANP, present in cardiac atrial tissue, has profound effects on salt and water homeostasis, as it acutely reduces blood volume by sequestering plasma and over the longer term by promoting renal salt and water excretion.

Energy Balance

Exogenous urocortin administration promotes negative energy balance, both by increasing energy expenditure and by decreasing food intake. The distribution of the urocortins supports the hypothesis that they may physiologically regulate energy balance via actions at CRF_2 receptors, exemplified by concordant hypothalamic expression in the VMH (Ucn 3), arcuate nucleus (Ucn 2, Ucn 3), and PVN (Ucn 1, Ucn 2) and in the NTS of the caudal hindbrain (Ucn 1). Furthermore, levels of VMH CRF_2 mRNA covary with circulating leptin levels, providing a homeostatic mechanism for modulating sensitivity to catabolic urocortin action. The presence of urocortins in the GI tract, glucose-regulating organs, and adipose tissue also is consistent with a possible peripheral role for urocortins in energy balance regulation.

With respect to energy expenditure, Ucn 1 (intracerebroventricular [ICV] administration) increases whole body oxygen consumption as measured by indirect calorimetry and increases sympathetic nervous system activity. The hypothalamus is a candidate site for metabolic action for urocortins, as intra-PVN Ucn 1 increases plasma leptin levels, induces intrascapular brown adipose tissue (BAT) uncoupling protein-1 (UCP1) mRNA synthesis, and increases the relative utilization of fat as an energy substrate. Whether energy expenditure-increasing effects of Ucn 1 are mediated in part by CRF_2 receptors and shared by the type 2 urocortins is unclear.

With respect to energy intake, peripheral Ucn 1 or stresscopin administration potently reduces food intake in rodents, possibly in part by slowing gastric emptying (see below). Central infusion of urocortins also potently suppresses feeding and reduces gastric emptying, effects that were shown to be at least partly CRF_2-mediated from studies that used selective agonists (type 2 urocortins), antagonists, antisense knockdown of receptor expression, and knockout mice. Unlike CRF_1 agonists, type 2 urocortins do not produce malaise, arousal, or anxiety-like effects at central anorectic doses in the rat. The degree to which endogenous brain urocortins produce these effects under physiological conditions remains unclear. In addition to the CRF_2-mediated urocortin anorexia, brain Ucn 1 may also reduce food intake via a separate CRF_1-dependent mechanism. Possible loci for Ucn 1-CRF_1-mediated anorexia include the dorsomedial nucleus of the hypothalamus, the parabrachial nucleus, and caudal hindbrain glucoregulatory sites.

Ucn 3 is present in pancreatic islet β cells, with pancreatic Ucn 3 levels responsive to changes in local glucose levels. Exogenous Ucn 3 administration increases glucagon and insulin levels *in vivo* and *in vitro*, resulting in a net increase in blood glucose levels. Glucoregulatory effects of Ucn 3 were abolished by pretreatment with a selective CRF_2 antagonist. Altogether, findings suggest an autocrine or paracrine regulation of glucose homeostasis by pancreatic Ucn 3.

Gastrocolonic Motility and Function

Stress also releases CRF-related peptides, possibly including urocortins, which slow gastric emptying through CRF_2 receptor systems. Through vagal efferents, central infusion of CRF receptor agonists reduce antral gastric motility, inhibit high-amplitude gastric contractions, and shift duodenal activity from fasted to fed motor patterns. Parallel to the central CRF_2 pathway for stress-induced gastric stasis, peripheral CRF_2 receptor activation delays gastric emptying, with peripheral (intravenous [IV] or intraperitoneal [IP]) administration of agonists with high (i.e., Ucn 1) or selective (i.e., Ucn 2 or Ucn 3) CRF_2 affinity delaying gastric emptying more potently than CRF. Like central infusion, peripheral Ucn 1 administration reduces antral gastric motility in fed rats and shifts gastric motor patterns from a fasted to fed state. Ucn 1 also hyperpolarizes stomach smooth muscle. Candidate regions that subserve CRF_2 receptor-mediated gastric stasis include, centrally, the PVN and dorsal vagal complex and, peripherally, myenteric fibers of the enteric nervous system. A physiological role for urocortins in stress-induced slowing of gastric emptying is supported by the presence of urocortins in the PVN, NTS of the caudal hindbrain, GI tract, and enteric nervous system.

Gastric Ucn 1 also may regulate gastric acid secretion, as peripheral administration of CRF_2 receptor agonists inhibits gastric acid secretion and increases gastric mucosal blood flow. Anatomical evidence supports such a relation. Ucn 1 colocalizes with tyrosine hydroxylase in parietal cells of the stomach, in proximity to CRF_2, but not CRF_1 receptors. CRF_2

receptors, in turn, colocalize with H /K -ATPase, the enzyme gastric proton pump, and somatostatin, which inhibits parietal cell activity and secretion of gastrin and histamine.

In contrast to their effects on gastric motor function, stressors stimulate colonic motility. Indicating a role for urocortin/CRF systems in these effects, central and peripheral administration of CRF_1, but not CRF_2, agonists likewise stimulates colonic motility, and selective CRF_1, but not CRF_2, antagonists attenuate stress-induced stimulation of colonic motility. Candidate substrates for CRF_1-mediated stimulation of colonic motor function include, centrally, the PVN and LC/Barrington nuclei that activate sacral parasympathetic nervous system activity and, peripherally, the colonic myenteric nervous system. Ucn 1 is expressed in these tissues, and, therefore, endogenous Ucn 1 may partly mediate the colonic-stimulating effects of stress.

Cardiovascular Function

Intravenous administration of urocortins more potently reduces blood pressure and has greater vasodilating effects than CRF. The relaxant effects of urocortin are mediated by arteriole and possibly cardiac CRF_2 receptors and appear to involve activation of protein kinase A and MAP kinase signaling pathways as well as inhibition of a PKC-dependent contractile mechanism. CRF_2-deficient mice have elevated mean arterial pressure, suggesting a physiological function of urocortin/CRF_2 interactions on vasculature *in vivo*. Independent of its effects on peripheral vascular resistance, systemic Ucn 1 infusion also had prolonged cardiac inotropic actions in sheep, increasing cardiac contractility, heart rate, and aortic blood flow.

Urocortins also are increasingly recognized to have cardioprotective and cardiovascular function-enhancing effects on compromised heart. Urocortins protect cardiac tissue from ischemia-reperfusion injury or hypoxia-reoxygenation injury via CRF_2 mechanisms. The cardioprotective effects of urocortins appear to involve activation of ERK1/2-p42, 44 signaling pathways, the phosphatidylinositol-3 (PI-3) kinase/Akt pathway, and protein kinase C epsilon isozyme activation. The cardioprotective effects also appear to involve reduced mitochondrial damage, achieved by increasing mitochondrial K_{ATP} channel opening. In addition to its cardioprotective effects, exogenous Ucn 1 also has cardiac hypertrophic effects, involving an Akt-dependent pathway.

Supporting a physiological cardioprotective role for urocortins, Ucn 1, 2, and 3 are present in the heart, with Ucn 2 and Ucn 3 abundant in myocardium. Plasma Ucn 1-IR increases in human systolic heart failure, and Ucn 1 expression increases more than ninefold in viable human myocytes after surgical cardioplegic arrest-reperfusion. Finally, cardiac myocytes isolated from CRF_2-deficient mice are more susceptible to ischemia-reperfusion injury.

Immune Function

As described previously, urocortins are expressed in immune tissue, including thymus, spleen, and/or skin, with Ucn 1 seen at the cellular level in lymphocytes, macrophages, fibroblasts, and mast cells. Exogenous Ucn 1 administration has palliative effects on experimental autoimmune encephalomyelitis and thermal injury-induced edema (although this may reflect glucocorticoid, vasodilatory, or direct cytoprotective mechanisms, as opposed to immunomodulating mechanisms). However, converging lines of evidence suggest that immune-derived urocortins locally modulate proinflammatory responses to perceived environmental insults.

For example, Ucn 1 mRNA is expressed in lamina propria macrophages, and Ucn 1-IR is detected throughout the entire lamina propria layer of the intestine. Intestinal Ucn 1 activity may be regulated by changes in the intestinal milieu, such as passage of dietary factors or food-associated bacterial antigens. Ucn 1-IR is elevated in the intestinal lamina propria macrophages of patients with ulcerative colitis, in which it is hypothesized to have a proinflammatory CRF_1-mediated effect. Intraperitoneal administration of CRF_1 agonists, such as Ucn 1, also increases intestinal mucosal permeability to macromolecules.

Other inflammatory conditions also are accompanied by increased Ucn 1 expression. In rheumatoid arthritis (but not osteoarthritis), Ucn 1-IR and Ucn 1 mRNA are elevated in synovium, and the number of Ucn 1-positive cells in synovia, including leukocytes and macrophages, is not only higher, but also correlates with the degree of inflammation. Supporting a proinflammatory property of secreted Ucn 1, Ucn 1 stimulates proinflammatory cytokine production by peripheral blood mononuclear cells *in vitro*, presumably, like CRF, via a CRF_1-mediated mechanism.

Some stress-related dermatological inflammatory conditions also have altered urocortin/CRF signaling. Mast cells, which play a role in allergy and inflammation, are widely distributed in the skin and synthesize and secrete both CRF and Ucn 1 in response to psychosocial stress or immunoglobulin E receptor cross-linking. Mast cells also express CRF receptors, activation of which leads to the release of proinflammatory mediators. Thus, atopic dermatitis, psoriasis, and alopecia areata have been

hypothesized to involve a stress-related precipitation or exacerbation of mast cell activation via local CRF/ Ucn 1–CRF receptor signaling.

The reproductive tract also contains mast cells. Ucn 1-IR is increased more than 10-fold in spontaneous abortion products from women with a history of multiple nonelective abortions relative to products from elective or nonhabitual abortions. Supporting the hypothesis that the observed increase in Ucn 1-IR is related to mast cell activation, levels of tryptase, which constitutes one-fifth of all protein in mast cells, and IL-8, an abortogenic mast cell-derived cytokine, also are robustly elevated in habitual spontaneous abortion products. Endometriosis also is associated with an increased number of activated mast cells in association with strong Ucn 1 immunostaining. Normal endometrium, in contrast, shows low tryptase and Ucn 1 immunoreactivity. Thus, Ucn 1 expression and mast cell activation are both elevated in inflammatory conditions of the reproductive tract and may correlate with increased risk for spontaneous abortion.

Finally, Ucn 1-IR is increased in stomach biopsies from patients with active gastritis. Unlike the colon, the stomach is much richer in CRF_2 than in CRF_1 receptors. As a result, in contrast to the CRF_1-mediated proinflammatory effects of intestinal Ucn 1, stomach Ucn 1 has been hypothesized, via CRF_2 receptors, to reduce/repair injury to gastric mucosa following noxious stimuli. Consistent with this perspective, Ucn 1-IR increases in treatment-responsive, but not treatment-resistant, patients during recovery from gastritis.

Reproductive Function

In addition to mast cells within the reproductive system, the human reproductive system itself expresses Ucn 1, CRF receptors, and CRF-BP. Ucn 1-IR is produced by choriodecidual and placental tissue and is elevated in maternal plasma from midgestation through birth. Placental Ucn 1 appears to act as a vasoactive relaxant factor on uteroplacental vasculature via local action at CRF receptors. Supporting the physiological relevance of this mechanism, pregnant women who have impaired uterine artery blood flow have significantly reduced circulating Ucn 1 levels in proportion to the degree of increased arterial resistance. Urocortins also may stimulate uterine contractility, as myometrium expresses both CRF_1 and CRF_2 receptors. Whereas CRF inhibits contractility via CRF_1 receptors, urocortins increase myometrial contractility via local CRF_2-mediated actions. Thus, urocortins may regulate placental vessel resistance to blood flow and stimulate uterine contractility.

Anxiety-Related Behavior

Central administration of Ucn 1, like CRF, has anxiogenic-like properties in many rodent anxiety models that are mediated at least partly by CRF_1 receptors. Ucn 1's effects on acoustic startle responding appear to differ from those of CRF, however, as exogenous Ucn 1 dampens rather than potentiates the acoustic startle response. Contrary to the effects of CRF_1 receptor agonists, ICV type 2 urocortin administration does not consistently increase anxiogenic-like behavior. For example, Ucn 3 did not increase anxiety-like behavior in the open field, elevated plus maze, light/dark box, social interaction, or defensive burying tests. Ucn 2 (ICV) similarly lacked acute anxiogenic-like effects in the rat open field and elevated plus maze tests. In fact, ICV Ucn 3 acutely produced anxiolytic-like changes in the elevated plus maze and light/dark box tests, and Ucn 2 had delayed, anxiolytic-like effects in the plus maze and reversed the anxiogenic-like effects of CRF in the open field. Similarly, whereas CRF_1 agonists increased activity in familiar environments, type 2 urocortins had mild motor suppressing effects.

However, stress-like effects of CRF_2 receptor activation also have been reported following local microinfusion into discrete brain regions, including the lateral septum and dorsal raphe, findings somewhat incongruous with the reviewed effects of ICV type 2 urocortin administration. Furthermore, unlike in rats, exogenous ICV administration of Ucn 2 to mice increased anxiety-like behavior in the plus maze as well as acoustic startle responses. Thus, although it is accepted that CRF_1 receptor activation has anxiogenic-like effects, general conclusions about the anxiety-related effects of central administration of the CRF_2-selective type 2 urocortins are not yet entirely possible and may involve brain region or species specificity. The endogenous roles of urocortins similarly remain unclear, as one Ucn 1-deficient mouse model exhibited normal anxiety-like behavior, whereas another showed increased anxiety-like responses.

Hearing

Ucn 1 may be required for normal hearing, as Ucn 1-deficient mutant mice have shorter hair cells in the inner spinal bundle of the organ of Corti and a higher response threshold in the auditory brain stem response examination than wild-type littermates. Perhaps reflecting this difference, a second murine model of Ucn 1 deficiency exhibited a reduced acoustic startle response. Anatomically supporting a relation of Ucn 1 to audition, Ucn 1-IR neurons are present in the lateral superior olive as well as in the inner spiral

bundle of the organ of Corti. In the organ of Corti, Ucn 1 appears at 8 days of age, present only in the inner hair cell region, in proximity to CRF_1 and CRF_2 receptors.

Acknowledgments

This work was supported by NIH grants DK70118, DK64871 and DK26741 from the National Institute of Diabetes and Digestive and Kidney Diseases. This is publication number 17400-NP from the Scripps Research Institute.

See Also the Following Articles

Anxiety; Blood Pressure; Cardiovascular System and Stress; Central Stress Neurocircuits; Corticotropin Releasing Factor (CRF); Corticotropin Releasing Factor-Binding Protein; Corticotropin-Releasing Factor Receptors; Cytokines; Gastrointestinal Effects; Immune Response; Immunity; Macrophages; Peptides; Psychosomatic Medicine; Somatic Disorders; Stress Effects, Overview; Ulceration, Gastric.

Further Reading

Bale, T. L. and Vale, W. W. (2004). CRF and CRF receptors: role in stress responsivity and other behaviors. *Annual Review of Pharmacology and Toxicology* **44**, 525–557.

Hashimoto, K., Nishiyama, M., Tanaka, Y., et al. (2004). Urocortins and corticotropin releasing factor type 2 receptors in the hypothalamus and the cardiovascular system. *Peptides* **25**, 1711–1721.

Scarabelli, T. and Knight, R. (2004). Urocortins: take them to heart. *Current Medicinal Chemistry: Cardiovascular and Hematological Agents* **2**, 335–342.

Suda, T., Kageyama, K., Sakihara, S., et al. (2004). Physiological roles of urocortins, human homologues of fish urotensin I, and their receptors. *Peptides* **25**, 1689–1701.

Theoharides, T. C., Donelan, J. M., Papadopoulou, N., et al. (2004). Mast cells as targets of corticotropin-releasing factor and related peptides. *Trends in Pharmacological Sciences* **25**, 563–568.

Zorrilla, E. P., Taché, Y. and Koob, G. F. (2003). Nibbling at CRF receptor control of feeding and gastrocolonic motility. *Trends in Pharmacological Sciences* **24**, 421–427.

V

Vaccination

V E Burns, A C Phillips and K M Edwards
University of Birmingham, Birmingham, UK

© 2007 Elsevier Inc. All rights reserved.

Introduction
The Vaccination Model
Stress and Antibody Response to Vaccination
Influence of Other Psychosocial Factors on Antibody Response to Vaccination
Possible Mechanisms
The Possible Paradoxical Effect of Acute Stress
Conclusion and Future Directions

Glossary

Antigen	Any molecule that generates and binds to an antibody.
B lymphocytes	A major type of lymphocyte; when activated by antigen, B lymphocytes differentiate into cells that produce antibodies against the antigen.
Immunoglobulin	Immunoglobulins are secreted by B lymphocytes and help fight infection through various mechanisms; also known as antibody.
In vitro techniques	Experimental techniques in which the experiment is performed outside a living organism; the conditions, and therefore the results, may not represent *in vivo* function.
In vivo techniques	Experimental techniques in which the research is conducted within the whole organism.
Neuroticism	A traitlike, or stable, personality dimension consisting of emotional instability and maladjustment.
Thymus-dependent antibody responses	Antibody responses, usually against a protein antigen, that require that the B lymphocytes gets help from T lymphocytes.
Thymus-independent antibody responses	Antibody responses, against a polysaccharide or lipopolysaccharide antigen, in which the B lymphocytes do not require help from T lymphocytes.
T lymphocytes	A major type of lymphocyte; one of their many functions is helping B lymphocytes produce an antibody response.
Vaccination	The injection of a dead or attenuated pathogen (disease-causing agent) in order to induce an adaptive immune response and protect against disease.

Introduction

Historically, the immune system has been considered to function in a relatively autonomous manner. Many years of research, however, have now revealed that the immune system is inextricably linked with the complex neuroendocrine networks of the body. These interactions provide plausible biological pathways through which psychological factors can influence immune function.

There is now considerable evidence that periods of chronic stress are associated with reductions in the number of a variety of immune cells in peripheral blood and with the impaired function of these cells *in vitro*. However, the clinical implications of these changes in healthy people, and within normal ranges, are unclear. Changes in cell number will not necessarily result in a significant change in the capacity of the lymphoid system to make an effective response to antigenic challenge; they may just reflect changes in the dynamics of lymphocyte migration and recirculation and other factors, such as shifts in plasma volume, rather than absolute changes in total cell numbers. Similarly, it is difficult to generalize *in vitro* measures of immune function, such as lymphocyte proliferation to mitogen and natural killer cell cytotoxicity, to *in vivo* processes. The isolated testing of any particular aspect of the network of immune cells provides limited information about the status of the highly integrated, complex immune system and, as such, has limited application to overall understanding of the relationship between stress and susceptibility to disease.

In contrast, the vaccination model affords an ecologically valid means of examining the associations

between psychosocial variables and the *in vivo* response of the immune system to antigenic challenge. Significantly, this response occurs within the functional context of the neuroendocrine system. A higher antibody response to vaccination is considered to be indicative, at a general level, of a well-functioning immune system and, at a more specific level, is likely to be associated with greater protection levels against the particular disease in question. As such, the adequacy of the immune response to vaccination has direct clinical relevance, both in terms of maximizing protection following vaccination and as a model of the response to naturalistic infection.

The Vaccination Model

Basic Immunology of Vaccination

The immune response to vaccination involves the coordination of a wide variety of immune cells. The vaccine antigen is initially recognized and presented by professional antigen-presenting cells, such as dendritic cells. The antigen is then recognized by specific T lymphocytes, which proliferate and differentiate into T-helper-2-cell lymphocytes. Vaccine antigen is also recognized by B lymphocytes without the necessity for antigen processing. When stimulated by antigen, B lymphocytes proliferate and mature into short-lived plasma cells, which produce the earliest immunoglobulin, IgM. In a primary response to a previously unseen antigen, peak IgM response occurs around 5 days after vaccination. The interaction between the activated T and B lymphocytes leads to the formation of germinal centers and, therefore, the production of high-affinity IgG and IgA antibody. This response is slower than the IgM response and reaches a peak around 28 days after vaccination. Secondary antibody responses, in which the immune system has been previously exposed to the antigen, are more rapid and of greater magnitude.

Types of Vaccination

The vaccination model also allows the examination of which aspects of the immune system may be vulnerable to psychosocial influence. The antibody response to most vaccines is thymus-dependent, in which T-lymphocyte help is required for the production of antibody by B lymphocytes. In contrast, for thymus-independent antibody responses against polysaccharide antigens, B lymphocytes generate an antibody response without T-cell help. If, for example, the effects of stress are restricted to antibody responses to thymus-dependent vaccines, this would imply that T lymphocytes are more susceptible to psychosocial influence than B lymphocytes.

Stress and Antibody Response to Vaccination

Cross-Sectional Studies

Early studies took advantage of the large number of people undergoing routine vaccination, such as meningitis C vaccination for all young adults and hepatitis B for medical students. The largest of these studies examined the association between life events stress and the final antibody titer against hepatitis B vaccination in two cohorts of students, vaccinated either in the past 12 months or at least 13 months previously. Whereas life events exposure was not related to antibody response in the recently vaccinated cohort, participants in the earlier vaccinated cohort who reported higher life events stress over the past year were over twice as likely to show an inadequate antibody titer as those in this cohort with lower life events exposure. This association withstood adjustment for variations in unhealthy behaviors (smoking, alcohol consumption, exercise, and sleep) and coping style. The results suggest that the immunogenicity of the hepatitis B vaccination may initially override any influence of psychological factors. Further, the larger number of individuals with inadequate antibody titers in the early vaccination cohort provides more power to detect effects. Nevertheless, this study implies that psychological stress may have its principal effects on the rate of deterioration of protection.

A similar study investigated the role of psychosocial variables in the response to meningitis C vaccination. This is a conjugate vaccination, in which a polysaccharide antigen is conjugated to a protein in order to generate a thymus-dependent antibody response. This is a common technique to improve the antibody response to thymus-independent antigens. Participants with high perceived stress and psychological distress levels were more likely to have low antibody levels against meningococcal C polysaccharide following the vaccination during the previous 16 months. This finding suggests that antibody responses to conjugate vaccination are also susceptible to psychosocial influences.

Prospective Studies in Elderly Populations

In the prospective studies of psychosocial influence on vaccination response, the experimenters administered the vaccines at controlled times, in order to be able to assess pre- as well as postantibody levels. This

is particularly important with common diseases, such as influenza, in which naturalistic exposure is common. In elderly populations, the caregiver-control model has been used to investigate the influence of the chronic stress of caring for a spouse with dementia on the antibody response to influenza vaccination. Two separate studies have demonstrated that spousal caregivers are less likely to achieve a fourfold increase in antibody titer at 1 month postvaccination than matched control participants. Because this is the clinical criterion of an adequate vaccination response, these effects are likely to result in increased susceptibility to infection.

One potential shortcoming of the caregiver-control model is that caregivers may differ from controls in a number of respects other than psychological stress, such as experiencing more physical strain. Significantly, it has been shown that former spousal caregivers display the same impaired antibody response to influenza vaccination as current caregivers. Although much of the physical demand of caregiving terminates with bereavement, it is reasonable to presume that the psychological stress continues. This argues that the poorer responses to vaccination observed in caregivers is the result of psychological stress rather than physical strain. Further evidence comes from a recent study in which the efficacy of the influenza vaccination was tested in a large group of elderly people living in the community. Individuals who had experienced bereavement in the previous year had significantly poorer antibody responses against one of the three influenza strains contained in the vaccination. This suggests that a variety of chronic stressors can have a detrimental effect on antibody response to the thymus-dependent influenza vaccination.

The one thymus-independent vaccination that has been examined in elderly adults is the vaccine against pneumococcus infection. Spousal caregivers were compared to former caregivers and age-matched controls. Although there were no group differences either 2 weeks or 1 month after the vaccination, the current caregivers had poorer specific antibody titers 3 and 6 months after receiving the pneumococcal vaccination than the other two groups. This suggests that caregiving impacts the rate of deterioration of protection rather than the initial response to thymus-independent vaccinations.

Prospective Studies in Young Populations

In younger populations, a common vaccine that has been used is the hepatitis B vaccination program. This consists of three separate vaccinations, administered at 0, 1, and 6 months. In one of the first studies in this area, individuals reporting higher mean perceived stress and anxiety over the whole vaccination period, probably indicating some trait or dispositional tendency, were less likely to have developed detectable antibodies against the vaccine by the time of the second inoculation. Although it is initially unclear why stress and anxiety levels assessed some months later may be related to this seroconversion, it is likely that these questionnaires measure, to some extent, trait or dispositional tendencies to report high stress and anxiety. It seems feasible, therefore, that people who generally tend to perceive their lives as stressful have higher mean stress scores and that it is these individuals who did not seroconvert at this time. It has also been demonstrated that the immunogenicity of the vaccine is important; a pair of studies by the same investigators found a negative effect of stress on antibody response to a low-dose hepatitis B vaccine, but the full-dose vaccination did not yield any significant stress effects.

Stress and the antibody response to the influenza vaccination has also been studied in younger adults. In one study, students who did not achieve a fourfold increase in antibody titer to the A-strains of the vaccine at 5 months postvaccination reported higher stressful life event exposures, and there was also a trend for those who were unprotected at 5 weeks to report more stressful life events in the period following vaccination. In a study of the effects of daily stress and feelings of being overwhelmed during the 10 days following vaccination, higher stress ratings were associated with lower antibody titers to one of the viral strains at both 1 and 4 months following vaccination. These studies provide evidence that stressful life events both preceding and in the period immediately following vaccination can influence the antibody response.

There has been only one study in which thymus-dependent and thymus-independent vaccines have been directly compared. This study examined the associations between stressful life events and the antibody response to the thymus-dependent influenza and thymus-independent meningococcal A + C vaccinations in a sample of 57 students. Psychosocial factors were associated with the antibody response to the influenza vaccination and were not associated with the response to thymus-independent meningococcal A polysaccharide vaccine component. This provides preliminary evidence that thymus-dependent vaccinations are more susceptible to the impact of stress than thymus-independent vaccinations.

Influence of Other Psychosocial Factors on Antibody Response to Vaccination

There is increasing interest in the influence of more stable psychological characteristics on antibody

response to vaccination. An early study in young girls found that higher internalizing, a concept similar to neuroticism, and lower self-esteem were associated with lower antibody titers following rubella vaccination. This was apparent, however, only in those who did not have serological evidence of exposure to rubella prior to the vaccination and were thus exhibiting a primary vaccine response. Trait negative affect has also been negatively associated with the secondary antibody response to the second hepatitis B injection. Similarly, neuroticism has been associated with a poorer antibody response to the influenza vaccine at both 5 weeks and 5 months postvaccination.

Social support has also been demonstrated to be associated with the response to vaccination. Two separate studies have demonstrated that students who reported greater social support have better antibody responses to either hepatitis B or influenza vaccination. Further, in a study of college freshmen, loneliness and smaller social network size were associated with a poorer antibody response to the influenza vaccination. In an elderly population, although generic social support was not associated with the influenza vaccination response, being married, and particularly reporting being in a happy marriage, was beneficial. With aging, marriage may become the primary social resource that influences health, including susceptibility to infection.

Possible Mechanisms

An association between stress and antibody response to vaccination could arise from direct and/or indirect mechanisms. The most likely indirect mechanisms are changes in health behaviors, often associated with periods of stress and their subsequent physiological impact on antibody formation. Poorer antibody responses to vaccination have been observed in those who consume high levels of alcohol, in regular smokers, and in those with inadequate nutrition. In spite of these findings, some studies have failed to adequately control for health behaviors, leaving open the possibility that at least some of the observed association between stress and antibody status may be attributable to variations in unhealthy behaviors.

More direct mechanisms of interaction between stress and immune function have also been postulated. Stress is associated with alterations in sympathetic nervous system and hypothalamic-pituitary-adrenal axis activation. Functional relationships between these neuroendocrine pathways and the immune system have been acknowledged for some time, including the innervation of lymphoid tissue and neuroendocrine receptors on many types of immune cells. Therefore, it is feasible that changes in the basal activity in these systems or their repeated activation in response to stress could impact on antibody development and maintenance.

The Possible Paradoxical Effect of Acute Stress

In contrast to the immunosuppressive effects of chronic stress discussed so far, it has also been suggested that acute stress may be immunoenhancing when experienced close to the immune challenge. Dhabhar and McEwen argued that immunoenhancement by acute stress is adaptive from an evolutionary stand point and might be regarded as an integral component of the fight-or-flight response. From this perspective, circumstances that initiate a fight-or-flight response are likely to also involve exposure to antigens and, therefore, a robust immune response would be adaptive for survival. There is now convincing evidence from animal studies of acute stressors in close temporal proximity to vaccination that provides support for an immunoenhancing effect of stress on the antibody response. Only one study has examined the effect of acute psychological stress on antibody response to vaccination in humans. Participants completed either a 45-min time-pressured, socially evaluated mental arithmetic task or a resting control period immediately prior to influenza vaccination. An enhancement of the antibody response to one of the influenza viral strains was found in women in the psychologically stressed group compared to controls.

Conclusion and Future Directions

There is now considerable evidence that chronic stress is associated with a poorer antibody response to a variety of vaccinations. These *in vivo* findings give greater credence to the often reported association between psychological factors and immune function. The next challenge is to examine whether interventions designed to reduce psychological stress or improve more positive psychological characteristics can result in enhanced vaccination responses. Further, it is important to determine whether the immunomodulation associated with either chronic or acute stress is sufficient to result in differences in actual infection rates following vaccination. These studies would more fully establish causality as well as having clear clinical implications for improving protection following vaccination, particularly in vulnerable populations.

See Also the Following Articles

Immune Cell Distribution, Effects of Stress on; Immune Function, Stress-Induced Enhancement; Immune Response; Immune Suppression; Immune System, Aging; Immunity; Depression, Immunological Aspects; Immune Surveillance – Cancer, Effects of Stress on.

Further Reading

Burns, V. (2004). Stress and antibody response to vaccination: implications of animal studies for human clinical research. *Expert Reviews in Vaccines* **3**, 141–149.

Burns, V. E., Carroll, D., Ring, C., et al. (2003). Antibody response to vaccination and psychosocial stress in humans: relationships and mechanisms. *Vaccine* **21**, 2523–2534.

Cohen, S., Miller, G. E. and Rabin, B. S. Psychological stress and antibody response to immunization: a critical review of the human literature. *Psychosomatic Medicine* **63**, 7–18.

Edwards, K. M., Burns, V. E., Reynolds, T., et al. (2006). Acute stress exposure prior to influenza vaccination enhances antibody response in women. *Brain, Behavior, and Immunity* **20**, 159–168.

Glaser, R., Kiecolt-Glaser, J. K., Malarkey, W., et al. (1998). The influence of psychological stress on the immune response to vaccines. *Annals of the New York Academy of Sciences* **840**, 649–655.

Phillips, A. C., Burns, V. E., Carroll, D., et al. (2005). The association between life events, social support, and antibody status following thymus-dependent and thymus-independent vaccinations in healthy young adults. *Brain, Behavior, and Immunity* **19**, 325–333.

Vedhara, K., Cox, N. K. M., Wilcock, G. K., et al. (1999). Chronic stress in elderly carers of dementia patients and antibody response to influenza. *Lancet* **353**, 627–631.

Vasoactive Peptides

W K Samson and M M White
Saint Louis University School of Medicine,
St. Louis, MO, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Z T Resch and W K Samson, volume 3, pp 643–649, © 2000, Elsevier Inc.

Vasoactive Peptides
Vasoactive Peptides and the Pituitary Response to Stress
Integration of Central and Peripheral Responses to Stress
A Question of Physiological Relevance

Glossary

Calcitonin gene-related peptide (CGRP)	A peptide homologous in structure and function to AM.
Chronotropism	The rate of cardiac muscle contraction.
Diuresis	The excretion of water in urine.
Inotropism	The strength of cardiac muscle contraction.
Intermedin	The recently discovered member of the adrenomedullin-calcitonin gene-related peptide (AM-CGRP) family of peptides.
Natriuresis	The excretion of sodium in urine.
Natriuretic peptides	A family of three homologous mammalian peptides, encoded by three distinct genes, that are produced in heart and the vasculature and that were initially discovered on the basis of their renal actions to stimulate sodium excretion.
Preproadrenomedullin-derived peptides	One mammalian gene encodes a prohormone that is posttranslationally modified to yield two vasoactive peptides: adrenomedullin (AM) and proadrenomedullin N-terminal 20 peptide (PAMP).
Receptor activity modifying proteins (RAMPs)	Receptor proteins that provide selectivity for the actions of adrenomedullin and calcitonin gene-related peptide.
Salt appetite	The behavior of selective ingestion of osmotically active solute.

Regardless of the operant definition, stress elicits significant, physiologically relevant changes in not only emotional behaviors, but also endocrine and cardiovascular function. Poised appropriately between the perception of a stressor and the integrative response to that insult are vasoactive peptides, factors that exert potent physiological effects in brain, pituitary gland, and the periphery. In many ways, these hormones serve to integrate the responses of multiple tissue systems in a manner parallel and complementary to the autonomic nervous system. Indeed, vasoactive peptides act in brain to modulate sympathetic efferent activity, fluid and electrolyte balance, and neuroendocrine responses to stress. In the pituitary gland, they significantly contribute to the regulation of

corticotroph function. In the periphery, these hormones determine vascular tone, cardiac pumping efficiency, adrenal hormone production and release, and, importantly, renal handling of fluid and electrolytes. Two families of vasoactive hormones, known to be released in response to numerous stressors, provide excellent examples of the integrative biology of the stress response. The natriuretic peptides and members of the adrenomedullin-calcitonin gene-related peptide (AM-CGRP) family of peptides are produced in brain and pituitary gland, as well as in the peripheral vasculature, heart, and kidney. They act both locally as autocrine factors and after release into the circulation as classic hormones to in some cases initiate and in others modify stress responses at the cellular and whole organ level.

Vasoactive Peptides

The simplest definition of a vasoactive peptide is a small protein that exerts either prorelaxant or procontractile effects in intact blood vessels. These effects may be endothelium dependent, requiring the recruitment of additional peptides, prostaglandins, or diffusable gases from the endothelial cell, or they may be endothelium independent, reflecting a direct effect on the vascular smooth muscle itself. Vascular tone is determined primarily by neurogenic mechanisms, largely mandated by the level of sympathetic activation. Circulating and locally produced factors also play important roles in the selective activation of distinct vascular beds and in the overall maintenance of vascular pressure, the single most important determinant of adequate perfusion of the vital organs in stress. Local factors not only directly affect vascular reactivity, but also, via presynaptic mechanisms, modify sympathetic tone. These factors also contribute to the maintenance of adequate perfusion pressures by altering cardiac performance and adjusting vascular volume.

Natriuretic Peptides

Originally discovered in extracts of cardiac atrium, the prototype of natriuretic peptides was originally named atrial natriuretic peptide (ANP) due to its isolation site, biological action, and protein structure. A second member of the family, originally thought to be produced in brain, but now known to be of almost exclusively myocyte (indeed, ventricular myocytes) origin, is B-type natriuretic peptide (BNP). Both ANP and BNP are released in response to volume overload, sensed by the myocardium as excess stretch. They share major actions to stimulate sodium and water excretion in the urine (Figure 1). ANP appears more relevant under physiological conditions, and BNP seems to be recruited in more chronic conditions of volume/pressure overload. Production of at least ANP also occurs in brain and pituitary gland, and actions there related to stress are numerous and diverse. The third member of this family of peptides, all products of unique genes, is C-type natriuretic peptide (CNP). Produced predominantly in brain and vasculature, very little, if any, CNP is produced in the myocardium. Although CNP is a poor natriuretic and diuretic agent, it is a profound vasodilator and can exert apparently paradoxical effects on sympathetic tone via central nervous system (CNS) actions. Much is known of the molecular biology, pharmacology, physiology, and pathophysiology of the natriuretic peptides.

Preproadrenomedullin-Derived Peptides and Their Homologs

Surprisingly, vasodilatory activities were isolated from extracts of a pheochromocytoma, and these peptides were later discovered to be posttranslational products of the same messenger RNA. Produced not only in adrenal medulla, but even more widely in the blood vessels, brain, and pituitary gland, these two vasodilatory factors exert complementary actions via unique signaling mechanisms (Figure 2). Adrenomedullin (AM), a 52-amino-acid peptide from the C-terminal portion of the proadrenomedullin precursor, exerts its vasorelaxant effects by stimulating nitric oxide synthase activity in endothelial cells of the vasculature. Proadrenomedullin N-terminal 20 peptide (PAMP), however, is a potent vasorelaxant due to its ability to exert presynaptic inhibition on sympathetic fibers innervating the vessel and to negatively modulate cholinergic activation of the chromaffin cells of the adrenal medulla and therefore reduce the level of epinephrine released in stress. Not all the actions of AM and PAMP are shared. For example, AM is a potent natriuretic and diuretic agent but does not affect catecholamine release significantly. A role for PAMP in sodium homeostasis also exists, however, as even in the absence of renotropic actions PAMP can stimulate sodium excretion due to its profound inhibitory effects on aldosterone secretion.

CGRP is a 37-amino-acid homolog of AM, and both are members of a larger family of peptides that includes calcitonin, amylin, and the recently described 47-amino-acid peptide intermedin (IMD; also named adrenomedullin-2). There is some evidence that these peptides may share an ancestral gene, and in addition to their structural similarities, they exert many shared biological activities (and perhaps receptors). CGRP, like AM, is a potent vasodilator when administered intravenously and has been demonstrated to inhibit aldosterone secretion *in vitro*.

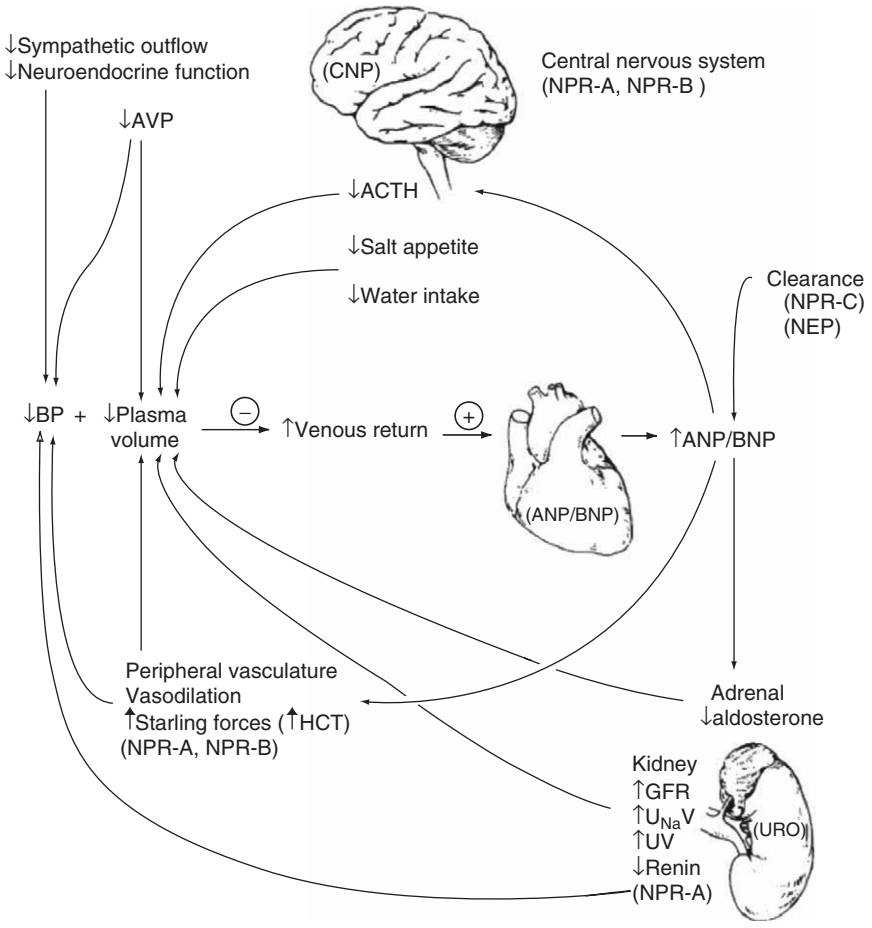


Figure 1 Pharmacological actions of the atrial natriuretic peptide (ANP) and B-type natriuretic peptide (BNP). AVP, vasopressin release; ACTH, adrenocorticotropin release; VR, venous return; BP, blood pressure, HCT, hematocrit; NPR-A, guanylyl cyclase A receptor; NPR-B, guanylyl cyclase B receptor; GFR, glomerular filtration; $U_{Na}V$, urinary sodium excretion; UV, urine volume. Also indicated is the production of C-type natriuretic peptide (CNP) in brain and urodilatin (URO), an N-terminally extended form of ANP in kidney. ANP and BNP are removed from the circulation by binding to the clearance receptor (NPR-C) and via degradation by neutral endopeptidase (NEP). Modified from the original and reprinted from Samson (1995), with permission of Humana Press.

When given into the CNS, like AM, it activates sympathetic tone, thus raising blood pressure. IMD exerts these same peripheral and central actions related to cardiovascular function and, like AM, exerts direct effects in the heart to increase contractility (inotropism) and beating frequency (chronotropism). Renal effects of IMD may parallel those of AM as well. It is not surprising that these homologous peptides exert similar actions (although unique effects may exist for each) since they share a common family of biological receptors. Those receptors are actually a complex of at least two proteins, the calcitonin receptor-like receptor (CRLR) and one of three possible receptor activity modifying proteins (RAMP1, RAMP2, or RAMP3). The relative combination of the respective RAMPs with the CRLR defines the specificity of the binding of each ligand. The CRLR–RAMP1 complex binds CGRP preferentially, while CRLRs in association with either RAMP2 or RAMP3 are the AM receptors. IMD binds promiscuously with all three CRLR–RAMP complexes, and thus it is not surprising that the hypotensive effects of IMD are at least partially blocked by relatively selective CGRP and/or AM antagonists.

Vasoactive Peptides and the Pituitary Response to Stress

Conventionally, the pituitary response to stress is considered to be the release of the hormone adrenocorticotropin (ACTH). The release of ACTH is controlled mainly by corticotropin-releasing hormone (CRH) levels in the hypophysial portal plasma; however, stimulation by vasopressin and oxytocin, as well as epinephrine, has been demonstrated. Thus, the effect of CRH can be modulated physiologically by factors also present in the portal circulation or by factors produced locally in the pituitary gland itself.

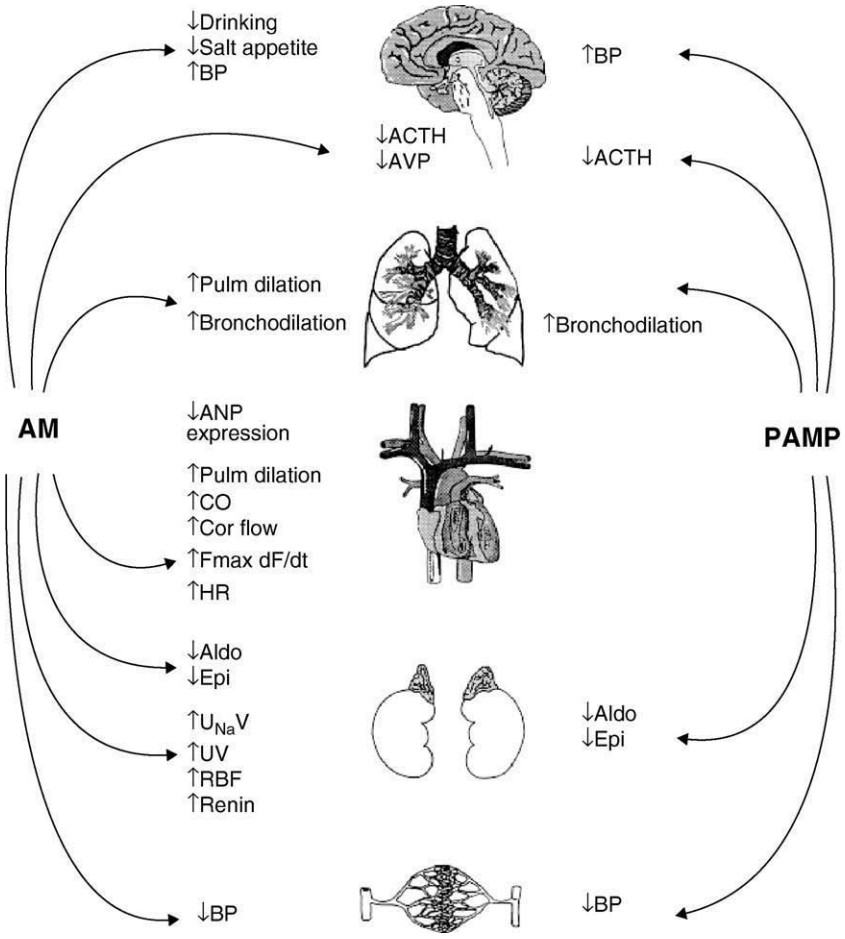


Figure 2 Pharmacological actions of adrenomedullin (AM) and proadrenomedullin N-terminal 20 peptide (PAMP). BP, blood pressure; ACTH, adrenocorticotropin secretion; AVP, vasopressin release; Pulm dilation, pulmonary dilation; ANP, atrial natriuretic peptide; CO, cardiac output, Cor Flow, coronary blood flow; Fmax dF/dt; maximum force of contraction; HR, heart rate; Aldo, aldosterone secretion, Epi, epinephrine release; $U_{Na}V$, urinary sodium excretion; UV, urine volume, RBF, renal blood flow, ET, endothelin production. Reproduced from Samson (1999), with permission from the *Annual Review of Physiology*, Volume 61, © 1999, by Annual Reviews.

Natriuretic Peptides

ANP is produced in the hypothalamus, where it exerts inhibitory effects on the releases of vasopressin and CRH. These effects may have physiological relevance, as stress-induced ACTH secretion is accentuated in rats pretreated centrally with anti-ANP antiserum, and glucocorticoid negative feedback is reduced when endogenous, brain-derived ANP is neutralized with the same antibodies. The effects are not limited only to ANP, as central administration of CNP also inhibits the stress response.

In addition to production in neurons projecting to stress centers in the hypothalamus (the paraventricular nucleus in particular), ANP and CNP are produced in cells of the anterior pituitary gland, where again CNP appears to be the most abundant gene product. Indeed, both the natriuretic peptide A receptor (NPR-A), which binds ANP preferentially, and the natriuretic peptide B receptor (NPR-B), which prefers CNP, are present. Even the third natriuretic peptide receptor (NPR-C), the clearance receptor, is found in the gland, although to a much lesser degree. Importantly, the NPR-A and NPR-B receptors have been localized to the corticotrophs, lactotrophs, and gonadotrophs. Both the NPR-A and the NPR-B are particulate guanylyl cyclase-linked receptors, activation of which results in elevated cellular levels of cGMP. ANP, but not CNP, potently stimulates the accumulation of cGMP levels in dispersed anterior pituitary cells. Activation of guanylyl cyclase is known in many cells to result in the stimulation of cGMP-dependent phoshodiesterases, which limit the activity of preformed cAMP. Because CRH stimulates ACTH secretion by activating adenylyl cyclase with the cAMP formed then stimulating protein kinase A activity, it is was not surprising that ANP, but not CNP, could inhibit CRH-induced ACTH secretion in cultured pituitary cells.

What pool of ANP could act in this fashion *in vivo*? ANP not only is produced in the pituitary gland, but also is delivered in the portal blood. Thus, the peptide could be acting to circumscribe ACTH secretion via a local paracrine mechanism or via a more classical neuroendocrine fashion. Are these pharmacological effects of ANP in cell culture physiologically relevant? While one group failed to observe any significant effects of physiological doses of ANP on CRH-stimulated ACTH secretion in humans, two other human trials did reveal a significant inhibitory effect of intravenously administered ANP on ACTH release.

Preproadrenomedullin-Derived Peptides

The adrenomedullin gene is transcribed in the hypothalamus, and AM-immunoreactive positive neurons projecting to neuroendocrine centers related to stress have been identified. Indeed, direct application of AM caused depolarization of parvocellular neurons of the paraventricular nucleus (PVN) in brain slices. Although the chemical phenotype of those AM-responsive parvocellular PVN neurons was not identified, they are composed of at least two classes of neurons, both playing a role in the hypothalamic control of the stress response. One subset projects to median eminence and delivers releasing factors, including CRH, to the hypophysial portal vessels for transport to the anterior pituitary gland. The other subset, also including some CRH-producing neurons, projects to autonomic centers in brain stem and thus may coordinate the cardiovascular stress response with that of the hypothalamic-pituitary-adrenocortical (HPA) axis. In fact, AM administration has been reported to increase Fos expression in CRH-positive hypothalamic neurons and to activate the HPA axis. Convincing evidence for a direct action of AM on CRH-producing neurons that project to median eminence comes from studies in which the ability of centrally administered AM to stimulate ACTH release was abrogated by pretreatment with a neutralizing anti-CRH antibody. The central action of CGRP to stimulate the HPA axis similarly was blocked by pretreatment with an anti-CRH antibody, and recently IMD was demonstrated to activate the HPA axis by stimulating CRH release into the median eminence.

While AM immunoreactivity is not abundant in the external layer of the median eminence, and thus a neuroendocrine role for AM or PAMP in the regulation of ACTH secretion cannot be predicted, AM gene products are widely distributed in the anterior lobe of the pituitary gland and binding for AM has been demonstrated. On the basis of the ability of AM and PAMP to inhibit angiotensin II and potassium-stimulated aldosterone secretion from zona glomerulosa cells *in vitro*, it was predicted that the peptides might also inhibit ACTH release from corticotrophs in culture. Basal and CRH-stimulated ACTH secretion *in vitro* is inhibited significantly by doses of AM, which approximate those present in the gland *in vivo*. The effects are dose and time dependent and are not related to the well-known ability of the peptide to activate adenylyl cyclase in peripheral tissues. In fact, the mechanism of the inhibitory effect of AM on ACTH remains unknown, as the peptide did not stimulate adenylyl or guanylyl cyclase activity in these cell cultures and similarly did not alter the levels of nitric oxide present during the incubations. A direct effect on the membrane polarization state could be hypothesized; however, a potassium channel blocker, glybenclamide, did not alter the action of AM to inhibit ACTH secretion under basal conditions.

The second posttranslational product of preproadrenomedullin (PAMP) also inhibits basal ACTH in a time- and dose-dependent fashion. It too exerts this effect at physiological concentrations and, like AM, does not directly alter adenylyl or guanylyl cyclase activity *in vitro*. Unlike AM, PAMP does not affect CRH-stimulated ACTH secretion. The ability of PAMP to reduce basal, but not stimulated, hormone secretion suggested an effect on the membrane polarization state. Because PAMP has been reported to act in other cell lines via an ATP-sensitive potassium channel, the pituitary effect was hypothesized to be due to opening of such a channel and hyperpolarization of the corticotroph membrane. Indeed, the ATP-sensitive potassium channel blocker glybenclamide completely blocked the inhibitory action of PAMP on ACTH secretion.

It remains to be determined whether these *in vitro* findings reflect physiological relevance in the *in vivo* setting. However, evidence for an *in vivo* effect of at least AM on ACTH secretion comes from studies in conscious sheep where the intravenous, but not intracerebroventricular, administration of AM lowered plasma levels of ACTH. It will be important to determine how the endogenous production of AM and PAMP in the pituitary gland is regulated. In the periphery, inflammatory stimuli such as interleukins and tumor necrosis factor-α are potent stimulators of AM gene transcription and peptide release. Thus, proinflammatory signals may also regulate pituitary production of the peptides as well, and proinflammatory cytokines are known effectors of ACTH secretion.

It is in the pituitary gland that the biological actions of the members of the AM-CGRP-IMD family of peptides diverge. Unlike the inhibitory effects of AM on ACTH release, CGRP has been reported to

stimulate hormone secretion via a cAMP-protein kinase A-dependent pathway. On the other hand, IMD has been reported to lack significant effects on basal or CRH-stimulated ACTH release. Furthermore, IMD inhibits releasing factor-stimulated growth hormone (GH) secretion *in vitro*, which is quite the opposite of the reported actions of AM and CGRP to stimulate GH secretion. The effect of IMD on somatotroph function suggests the existence of an IMD receptor unique from the CRLR-RAMP receptor complexes. This *in vitro* action of IMD may be physiologically relevant since in rodents stress results in inhibition of GH release and this peptide acts both centrally and at the level of the pituitary gland to decrease GH release. The actual source of the IMD that may be exerting these actions is unknown; however, the gastrointestinal tract is a prime candidate, since the highest levels of the peptide have been detected in stomach, as well as in kidney and brain.

All three members of the AM-CGRP-IMD family of peptides exert behavioral effects that may reflect a response to stress. The peptides inhibit food appetite and water drinking when administered into the brain, and at least in the case of AM, the effects on water drinking and salt appetite have been shown to be physiologically relevant. It appears that the actions of AM are more related to fluid and electrolyte homeostasis and those of CGRP to energy intake. Intermedin exerts anorexigenic effects similar to those of CGRP; however, the actions of IMD on thirst more closely parallel those of AM.

Integration of Central and Peripheral Responses to Stress

The integrative response to stress requires a coordination of multiple organ systems. Because of their diverse sites and mechanisms of action, vasoactive peptides are excellent candidates for physiological regulators of these integrative responses. Natriuretic peptides provide the best example of factors that exert complementary actions in multiple tissue sites. Their ability to unload excess vascular volume via direct natriuretic and diuretic actions in kidney is matched by inhibitory actions on releases of the volume conservatory hormones vasopressin, renin, and aldosterone. Behavioral actions within brain also complement these renal effects. ANP not only stimulates fluid and electrolyte excretion, but also, in what is now known to be a physiologically relevant manner, inhibits water drinking and salt intake, thus allowing time for the renal mechanisms to compensate before the addition of new fluid volume to the already taxed vascular tree. In states of volume depletion, these mechanisms are inhibited, thus permitting sodium and water retention and additional intakes to replenish lost volume. The blood pressure-lowering effects of the natriuretic peptides are not remarkable, but they can be demonstrated at pharmacological doses *in vivo*, and data from ANP and NPR-A knockout mice suggest an important role for the peptide in blood pressure regulation.

In times of excess secretion, such as during cardiac overload states, endogenous natriuretic peptides may unload the heart by a vasodilatory mechanism that is complemented by increases in vascular permeability. One group has demonstrated a sympatholytic effect of ANP in the brain stem, which would also lower vascular pressure. Thus, the actions of natriuretic peptides seem well coordinated to facilitate fluid and electrolyte balance and the response to volume stress. Whether the peptides play any role in psychogenic stress remains to be seen. However, the ability of both ANP and CNP to inhibit the hypothalamic drive for gonadotropin secretion suggests that some stress-induced infertility may involve the hypothalamic actions of these peptides. Evidence for the physiological relevance of this effect has been provided in studies employing selective cell targeting of NPR-B receptive neurons in hypothalamus.

Adrenomedullin gene products provide a very different example of the potential importance of vasoactive peptides in the body's response to stress. Because of its profound effects on brain, pituitary, vascular, adrenal, and renal mechanisms related to fluid and electrolyte homeostasis, overexposure of the body to high levels of AM could have disastrous consequences. AM potently inhibits water drinking and sodium appetite, ACTH secretion, aldosterone release, vascular tone, and renal fluid reabsorption. Thus, when released in large amounts, as in sepsis, potentially fatal hypotension and the resultant underperfusion of vital organ systems could result. Fortunately, the actions of AM, and to some degree PAMP, include compensatory effects that protect against cardiac collapse. Indeed, in clinical studies the levels of AM attained in plasma during sepsis predict survival, probably due to the cardioprotective effects of those high levels of peptide. Furthermore, transgenic overexpression of the AM gene in mice results in enhanced survival outcomes during experimentally induced septic shock.

How is this accomplished? In brain, AM and PAMP are sympathostimulatory. This may seem a paradox in light of the profound hypotensive effects of the peptide in the periphery. However, their action in brain may compensate for not only the vasorelaxant effects of the peptides, but also the ability of AM to dampen baroreflexes and thus the perception of volume depletion. This central activation of

sympathetic efferents may assure adequate perfusion of vital tissues. Similarly, direct effects of AM on cardiac contractility have been documented. While decreasing venous return by peripheral vasodilation, AM also exerts positive inotropic and chronotropic effects in heart, which can maintain ejection fraction and stroke volume, again protecting against cardiac collapse. At least one additional protective effect of AM has been demonstrated. Even in the face of profound hypotension, renal perfusion pressure is maintained due to the dilatory effects of the peptide on both afferent and efferent renal arterioles. Thus, AM and PAMP exert protective effects that assure perfusion of the vital organ systems during hypotension, and these mechanisms must be considered important for the response of an individual to stress.

A Question of Physiological Relevance

With the exceptions of CNS actions of both ANP and AM to inhibit thirst and salt appetite, of AM to stimulate vasopressin release, of CNP to inhibit GnRH release, and of ANP to contribute to blood pressure regulation during sodium excess, much of the current knowledge of the physiological relevance of the vasoactive peptides remains based solely on pharmacological evidences. Only recently have combinations of methodologies been attempted to identify and confirm physiological relevance. For example, the salt appetite-suppressive effect of AM was suggested to have physiological relevance by studies employing selective antibodies that neutralized endogenous brain-derived AM. These initial findings were then confirmed by the antisense oligonucleotide approach in studies in which excessive salt appetite occurred when the endogenous production of AM was interrupted.

Why is it important to establish the physiological relevance of any peptide's pharmacological effects? The obvious answer is the potential use of an exogenous peptide in a therapeutic setting. Less obvious is the importance of elucidating the potential involvement of a given peptide in the genesis of a pathological state. Because of their pleuripotent biological activities, vasoactive peptides hold promise for both therapeutic use and the diagnosis of evolving pathology. In particular, plasma BNP levels are now accepted as diagnostic of the progression of cardiac overload in congestive heart failure in humans, and ANP is a clinical tool employed to unload excessive vascular volume in this disease. The ACTH-inhibiting effect of ANP during such treatment would certainly provide an extra therapeutic advantage. This again points to the significance of multiple sites and mechanisms of actions of vasoactive peptides during challenges to normal homeostasis, which by any definition mean stress.

See Also the Following Articles

Hypertension; Peptides; Salt Appetite.

Further Reading

John, S. W. M., Krege, J. H., Oliver, P., et al. (1995). Genetically decreased levels of atrial natriuretic peptide and salt-sensitive hypertension. *Science* **267**, 679–681.

Samson, W. K. (1995). Cardiovascular hormones. In: Conn, P. M. & Melmed, S. (eds.) *Endocrinology: basic and clinical principles*, pp. 361–376. Totowa, NJ: Humana Press.

Samson, W. K. and Taylor, M. M. (2005). Cardiovascular hormones. In: Conn, P. M. & Melmed, S. (eds.) *Endocrinology: basic and clinical principles* (2nd edn., pp. 321–335). Totowa, NJ: Humana Press.

Taylor, M. M. and Samson, W. K. (2002). Adrenomedullin and the integrative physiology of fluid and electrolyte balance. *Microscopy Research and Techniques* **57**, 105–109.

Taylor, M. M., Bagley, S. L. and Samson, W. K. (2004). A possible mechanism for the action of adrenomedullin in brain to stimulate stress hormone secretion. *Endocrinology* **145**, 4890–4896.

Taylor, M. M., Bagley, S. L. and Samson, W. K. (2005). Intermedin/adrenomedullin-2 acts within central nervous system to elevate blood pressure and inhibit food and water intake. *American Journal of Physiology* **288**, R919–R927.

Taylor, M. M., Bagley, S. L. and Samson, W. K. (2005). Stress hormone secretion is altered by central administration of intermedin/adrenomedullin-2. *Brain Research* **1045**, 199–205.

Taylor, M. M., Bagley, S. L. and Samson, W. K. (2006). Intermedin/adrenomedullin-2 inhibits growth hormone release from cultured, primary anterior pituitary cells. *Endocrinology* **147**, 859–864.

Vasopressin

L P Renaud
Ottawa Health Research Institute and
University of Ottawa, Ottawa, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by L P Renaud, volume 3, pp 650–655, © 2000, Elsevier Inc.

The Posterior Pituitary Hormones
Vasopressin Synthesis
Vasopressin Release
Receptors
Vasopressin in the Brain
Vasopressin and the Hypothalamic-Pituitary Axis

Glossary

Diabetes insipidus	A condition caused by insufficient vasopressin in the circulation or inability of vasopressin to activate receptors in the kidney; manifest by polyuria and polydipsia.
Hormone	A peptide that circulates in the blood and exerts its biological effects by binding to specific receptors on select target tissues.
Hypothalamus	The part of the central nervous system residing above the pituitary, containing neurons that control vital functions (e.g., pituitary secretions, thermoregulation, and homeostasis in the autonomic, cardiovascular, and hydromineral systems).
Median eminence	The area at the base of the third cerebral ventricle that contains two zones: an inner zone that has axons in passage to the posterior pituitary and an external zone that contains the axon terminals of parvocellular neurons that synthesize and release peptides into adjacent portal capillaries.
Neurohypophysis	The posterior part of the pituitary, composed mainly of axon terminals of hypothalamic magnocellular neurosecretory neurons.
Nonapeptides	Peptides made up of nine amino acids.

Vasopressin is a nonapeptide that functions as a hormone in the blood and as a neurotransmitter or neuromodulator in the brain. Vasopressin is synthesized as a precursor molecule, prepropressophysin, within neurons with specialized functions. In magnocellular neurosecretory neurons of the hypothalamus, the precursor is stored in vesicles and undergoes posttranslational processing during axonal transport to terminals in the neurohypophysis. Vasopressin release into the systemic circulation occurs in response to the activation of these neurons by specific stressors that challenge hydromineral balance (e.g., dehydration) and cardiovascular homeostasis (e.g., blood loss and hypotension). Once in the circulation, vasopressin assumes key hormonal roles by binding to V_1 type receptors on vascular smooth muscle to produce vasoconstriction and to V_2 type receptors in the kidney to increase water reabsorption and antidiuresis. Other peripheral actions include platelet aggregation; liver glycogenolysis; and secretion of aldosterone, insulin, and atrial natriuretic factor by the adrenal gland, pancreas, and heart, respectively. Vasopressin synthesized centrally in other hypothalamic and extra-hypothalamic neurons in the brain acts mainly through V_1 type receptors to mediate several functions, including the modulation of neuronal excitability via autocrine, paracrine, or transmitter mechanisms, to augment the release of adrenocorticotropic hormone (ACTH) in the anterior pituitary during stress, to attenuate febrile responses, and to influence cognition and complex social behaviors.

The Posterior Pituitary Hormones

The discovery in 1895 that injections of extracts of beef pituitary increased arterial blood pressure marked the beginning of a century of research into the biology of the neurohypophysial hormones, identified as two structurally similar nonapeptides, oxytocin and vasopressin. Their main biological effects, the contraction of vascular and uterine smooth muscle and antidiuresis, were initially assessed by sensitive bioassays. Elucidation in the 1950s of the chemical structure of vasopressin as distinct from oxytocin permitted the development of specific radioimmunoasseys, receptor agonists and antagonists, and techniques to investigate the molecular biology and function of vasopressin and oxytocin genes. Although discovered through their peripheral hormonal actions following their release from the posterior pituitary, an important aspect of this evolution was the recognition that vasopressin and oxytocin are also neuropeptides whose synthesis and release by different populations of neurons in the brain can influence a variety of central nervous system functions.

Vasopressin Synthesis

In Hypothalamic Magnocellular Neurons

Vasopressin and oxytocin molecules circulating as hormones in the blood originate from specialized magnocellular neurons, termed neurosecretory cells, located within the hypothalamic supraoptic and paraventricular nuclei and accessory magnocellular nuclei. The accessibility of the magnocellular neurohypophysial neuronal system has provided investigators with a valuable model to study peptide biosynthesis. In accordance with their separate biological roles, the genes for vasopressin and oxytocin are usually localized in different magnocellular neurons. Vasopressin- and oxytocin-synthesizing neurons also contain genes that code for other neuropeptides, and these may be co-activated and co-expressed together with vasopressin (or oxytocin) under appropriate circumstances. Consistent with other neuropeptides, synthesis occurs in the cell somata, initially resulting in the formation of a large precursor molecule called prepropressophysin that is composed of a signal peptide, the nonapeptide vasopressin (arginine-8 vasopressin, AVP, in most mammals; lysine-8 vasopressin in pig), a putative carrier molecule neurophysin, and a 39-amino-acid glyocoprotein (co-peptin, CCP). After packaging into large (160-nm) neurosecretory vesicles, prepropressophysin undergoes further posttranslational processing during transport in axons to storage sites in the posterior pituitary.

Magnocellular neurosecretory neurons constitutively express vasopressin in their somata and axons. They also demonstrate an increase in vasopressin hnRNA and mRNA within minutes of exposure to a physiological stimulus (examples include a rise in plasma osmotic pressure and the unloading of peripheral baroreceptors as a consequence of blood loss). Osmotic, hypotensive, and immune stimuli are also powerful activators of several immediate early genes, including cFos and cJun, in these neurons. Details of the linkage between these genes and the signal transduction mechanisms that regulate vasopressin gene transcription in health and disease (e.g., ectopic production by small-cell lung cancers) are topics under investigation.

A lack of vasopressin in animals and humans results in diabetes insipidus, a condition characterized by polyuria and polydipsia. Central diabetes insipidus may occur because of a lesion in the hypothalamic-neurohypophysial pathway or, more rarely, as an autosomal dominant inheritance in which mutations have been detected in the vasopressin signal peptide or neurophysin molecules. The only genetic animal model for hypothalamic diabetes insipidus is the Brattleboro rat; discovered 30 years ago, this animal fails to produce sufficient vasopressin due to a single base pair deletion in its vasopressin gene. The recent development of a vasopressin-null mouse may provide new insights into various roles of vasopressin in the stress response, cardiovascular regulation, and behavior.

In Parvocellular Neurons

Immunocytochemical and radioimmunoassay techniques reveal that vasopressin is present in other areas of the brain and spinal cord and in certain peripheral autonomic ganglia. In the brain, distinct groups of vasopressin-synthesizing parvocellular neurons may be recognized by their vasopressin-like immunoreactivity or through their ability to express vasopressin mRNA and hnRNA. The latter forms of gene expression reveal that vasopressin synthesis may depend on specific stimuli. For example, in the hypothalamic paraventricular nucleus, corticotropin releasing hormone (CRH)-synthesizing neurons may co-express vasopressin only during stress.

Vasopressin Release

Into the Systemic Circulation

Levels of vasopressin in the circulation (expressed as picograms per milliliter) are normally low to undetectable. Vasopressin is released into the systemic circulation from storage sites in the neurohypophysis in response to specific challenges to cardiovascular or hydromineral homeostasis, causing plasma levels to rise to >20 pg/ml. In some clinical circumstances associated with head trauma, congestive heart failure, liver cirrhosis, and cancer (e.g., small-cell lung carcinomas), humans have been found to have an excess of vasopressin in the circulation relative to their plasma osmolality, resulting in free-water retention and hyponatremia, a syndrome of inappropriate antidiuresis.

Circulating levels of the hormone are normally regulated through release from the neurohypophysis, a structure comprising mainly pituicytes and axon terminals of the hypothalamic magnocellular neurons that contain stored neurosecretory vesicles. Action potentials generated at the cell somata of the magnocellular neurons (in response to specific stimuli) propagate into the nerve terminals, thereby activating voltage-gated calcium channels that promote a rise in intracellular calcium and triggering exocytosis. The efficiency of this stimulus–secretion coupling mechanism is dependent on both the rate and pattern of discharge of action potentials invading the nerve terminals; release is optimized not only by an increase

in action potential frequency up to and beyond 20 Hz but also by the emergence of bursting or phasic patterns of firing.

Stimuli for release

Plasma hyperosmolality The osmolality of the plasma is sensed in the brain by a central osmosensor complex composed of neurons in the lamina terminalis along the anterior wall of the third cerebral ventricle in addition to the hypothalamic magnocellular neurons themselves. Recent electrophysiological studies indicate that magnocellular neurons are intrinsically osmosensitive; they contain mechanosensitive stretch-inactivated cation channels that sense changes in cell volumes consequent to a deviation from resting plasma osmolality (~290 mosmol/kg in humans). A rise in the osmolality of the extracellular environment initiates cell shrinkage; an ensuing increase in a membrane cationic conductance results in membrane depolarization and a subsequent increase in action potential generation once the membrane potential reaches threshold (around –50 mV). In the rat, the increase in activity of vasopressin-synthesizing magnocellular neurons usually initiates a unique phasic-bursting pattern of firing, an efficient trigger for stimulus–secretion coupling.

Hypovolemia and hypotension In mammals, information about the state of the peripheral cardiovascular system is continuously supplied to the brain via afferent pathways in the glossopharyngeal and vagus cranial nerves. Baroreceptors monitor the tone in the main arteries around the heart and lungs. A drop in arterial blood pressure, as might result from hemorrhage, decreases the input from baroreceptors. This information is channeled through catecholaminergic pathways that ascend through the brain stem to the hypothalamic magnocellular neurons, triggering an increase in their activity and initiating hormone release in the posterior pituitary.

In the Pituitary Portal Circulation

The external zone of the median eminence contains an abundance of axon terminals. Those that demonstrate vasopressin-like immunoreactivity arise mainly from parvocellular neurons in the hypothalamic paraventricular nucleus. A subpopulation of these are neurons that normally synthesize only CRH but that also contain inactive vasopressin genes that become expressed under conditions of prolonged stress and then co-release both peptides into the pituitary portal circulation from vesicles stored in these median eminence axon terminals. Receptors for vasopressin reside on anterior pituitary corticotrophs.

In the Brain

The neurohypophysial peptides are found in relatively high concentrations (in the picogram per microliter range) within the extracellular space in specific areas of brain. Because blood-born peptides are not seen in these concentrations and do not cross the blood–brain barrier in significant amounts, this reflects central release. Two likely sources are synaptic release from the axon terminals of central vasopressin-synthesizing neurons and exocytosis from somata and dendrites. The latter event is now well recognized as a mode of peptide release in the hypothalamic magnocellular neurons of the paraventricular and supraoptic nuclei. Here, high concentrations of vasopressin and/or oxytocin can be measured by microdialysis, cell somata and dendrites display neurosecretory vesicles with exocytotic profiles, and electrical events consistent with dendritic release can be recorded.

Stimuli for brain release Stressors such as social defeat and forced swimming are potent stimuli to trigger the release of vasopressin into the brain. Release sites may vary with the stimulus. For instance, social defeat is associated with vasopressin release in the paraventricular nucleus but not in the supraoptic nucleus and causes no change in levels of vasopressin in the plasma. Forced swimming, on the other hand, promotes vasopressin release in both the paraventricular and supraoptic nuclei, again with virtually no change in plasma vasopressin levels. This illustrates a dissociation between the mechanisms that trigger central versus peripheral release. Central release mechanisms have not been studied as extensively as those in the posterior pituitary.

In the Cerebrospinal Fluid

Vasopressin is present in the cerebrospinal fluid of many mammals. Unlike levels in the plasma, cerebrospinal vasopressin may exhibit large daily rhythms with high concentrations during daylight hours, suggesting that the origin is regulated by the suprachiasmatic nucleus, a regulator of circadian rhythms in the mammalian brain.

Receptors

Peripheral and central effects of AVP are mediated by at least three subtypes of G-protein-linked membrane-bound receptors. These fall into two classes, V_1 (V_{1a} vascular and V_{1b}/V_3 pituitary) and V_2, based on second-messenger cascades and pharmacological properties. The activation of V_1 class of receptor

subtypes ($V_{1a}R$ and $V_{1b}R$) results in the hydrolysis of phosphatidyl inositol and an increase in cytosolic calcium. The activation of the V_2 receptor (V_2R) results in the stimulation of adenylate cyclase.

The V_1 Receptors

This receptor has been cloned, and its mRNA (visualized by *in situ* hybridization histochemistry) is seen throughout the brain, liver, and kidney; the $V_{1a}R$ mediates the vasoconstrictor and hepatic glycogenolytic actions of vasopressin and is thought to be the predominant vasopressin receptor in brain. A subtype of the V_1 class, the $V_{1b}R/V_3R$, has been characterized in the anterior pituitary where it regulates vasopressin-mediated ACTH release by potentiating the actions of CRH. $V_{1b}R$ mRNA has also been detected in the thymus, heart, lung, spleen, kidney, uterus, breast, and brain.

V_2 Receptors

The V_2 receptors and V_2R mRNA are abundant in the kidney renal tubules, where they mediate the well-established antidiuretic properties by promoting free-water reabsorption. Aquaporin 2, the water channel in apical collecting duct cells, is present in the membrane of cytoplasmic vesicles and is translocated to the apical membrane in response to vasopressin. Vasopressin regulates two types of aquaporin 2 expression: one prompt and the other sustained. The failure of vasopressin action in the kidney renal tubules results in nephrogenic diabetes insipidus, a disorder that may be inherited, resulting from mutations in the genes for vasopressin V_2 receptors and/or aquaporin 2.

Vasopressin in the Brain

Initial studies on avoidance behavior in the early 1960s provided the first clue that vasopressin and related peptides were involved in learning and memory processes. It is now recognized that centrally released vasopressin is involved in a variety of roles, including rewarded behavior and drug tolerance; social, reproductive, and feeding behaviors; temperature and cardiovascular regulation; stress; and circadian rhythms.

Vasopressin Pathways

Techniques that can specifically define vasopressin-relevant regions of brain (ligand-binding assays, immunocytochemistry, and *in situ* hybridization) in combination with lesions and tract-tracing methods serve to define a chemical neuroanatomy of vasopressin in the brain. Prominent among these are the magnocellular neurons of the hypothalamic supraoptic, paraventricular, and accessory magnocellular nuclei and cells groups whose axons project to the posterior pituitary. The paraventricular nuclei also contain subsets of parvocellular neurons that synthesize vasopressin. Some of these neurons project to the median eminence and form a component of the hypothalamic-pituitary-adrenal axis; the expression of their vasopressin genes may be evident only under certain circumstances.

Other paraventricular and hypothalamic neurons that contain vasopressin constitutively project axons to the parabrachial nucleus, dorsal vagal complex, ambiguus nucleus, medial and ventrolateral medulla, and spinal cord areas; these are likely to participate in control of regulatory functions attributed to the autonomic, sympathetic, and parasympathetic nervous systems (e.g., heart rate, blood pressure regulation, and release of adrenal medullary hormones).

Vasopressin-immunoreactive neurons occupy the dorsomedial core of the hypothalamic suprachiasmatic nucleus (SCN), site of a biological clock in the mammalian brain. These cells project into the medial hypothalamus and contribute to the entrainment of circadian rhythmicity in hypothalamic neuroendocrine and thermoregulatory functions. Functions of a prominent extrahypothalamic projection to the midline thalamic paraventricular nucleus remain to be defined. SCN vasopressin-synthesizing neurons are the likely source for the circadian rhythmicity in vasopressin observed in the cerebrospinal fluid.

Vasopressin neurons in the bed nucleus of the stria terminalis and medial amygdaloid nucleus project to the lateral septum, ventral septal nucleus, olfactory tubercle, lateral habenular nucleus, midbrain central gray, dorsal raphe nucleus, locus coeruleus, and ventral hippocampus.

Factors Affecting Central Vasopressin Gene Expression

Differing levels of mRNA indicate that vasopressin expression is not uniform. For example, adrenal steroids block vasopressin mRNA expression in the parvocellular neurons in the paraventricular nucleus but not in other brain areas. Vasopressin expression in lateral septum and medial amygdala requires the presence of gonadal steroids; most vasopressin mRNA disappears within a week of castration. How gonadal steroids work remains unclear because there are no recognizable androgen- and estrogen-responsive elements on the promotor region of the vasopressin gene. Presumably, these steroids do not affect mRNA transcription but, rather, regulate vasopressin mRNA levels at a posttranscriptional level because castration decreases the length of the mRNA poly(A) tail in the

bed nucleus of the stria terminalis, whereas testosterone increases it, thereby influencing the stability of the vasopressin mRNA.

Electrophysiology of Central Vasopressin Receptors

The exogenous application of vasopressin alters the excitability of neurons in a variety of areas, including the hippocampus, hypothalamus, brain stem, and spinal cord, all regions containing vasopressin fibers and/or neurons. In most instances, the response to vasopressin is manifested as an increase in neuronal firing, blockable with V_1 receptor antagonists. Mechanisms vary with the areas tested. Most investigations have focused on direct, tetrodotoxin-resistant actions mediated via postsynaptic receptors. For example, in rat facial nucleus, vasopressin receptors produce depolarization mediated by nonselective cationic channels. In the lateral horn of the neonatal rat spinal cord, vasopressin acting at G-protein-coupled V_1 type receptors produces depolarization in spinal preganglionic and other lateral and ventral horn neurons by decreasing a resting potassium ionic conductance and/or increasing a nonselective cationic conductance. Other actions are indirect, involving presynaptic receptors that attenuate glutamatergic or GABAergic neurotransmission. As previously reported, vasopressin is released locally by magnocellular neurons and the peptide is proposed to act in an autocrine or paracrine function to alter the excitability of the neurons responsible for its synthesis, either directly or through synaptic modulation. Calcium-imaging studies in dissociated vasopressin-synthesizing neurons from rat supraoptic nucleus reveal that signaling at vasopressin receptors involves calcium influx through L, N, and T type channels, an autoregulation that seems to involve the activation of both protein kinase C- and adenylate cyclase-linked intracellular pathways.

Vasopressin and Thermoregulation

Studies in a variety of mammals have demonstrated the fever-reducing effects of vasopressin and the fever-enhancing actions of V_{1a} receptor antagonists injected into the ventral septal area. It is likely that this area receives its innervation from the bed nucleus of the stria terminalis. Interestingly, this is where electrical stimulation can attenuate pyrogen-induced fever, an effect that is blockable by injection of a V_{1a} receptor antagonist into the ventral septal area. Castration, which is followed by the disappearance of vasopressin in the septal area, lengthens pyrogen-induced fever, presumably because of a reduction in the amount of releasable vasopressin.

Vasopressin and Cognition

Extensive animal research has established a role for vasopressin in learning and memory. Less clear are the outcomes of similar efforts in humans. The hippocampal system has a prominent role in learning and memory, and current observations indicate a facilitating effect of vasopressin and vasopressin analogs on retrieval and relearning processes, mediated by both vasopressin V_1 type and oxytocin receptors located mainly in the ventral hippocampus.

Vasopressin, Monogamy, and Social Behavior

The neurohypophysial peptides have been implicated in the central mediation of complex social behaviors such as affiliation, parental care, and territorial aggression. For example, in the prairie vole, research suggests that vasopressin is involved in the control of several behaviors associated with monogamy, including pair bonding, paternal care, and mate guarding. Molecular studies suggest that changes in the regulation of oxytocin- and vasopressin-receptor gene expression underlie species differences in receptor distribution and might provide a mechanism for the evolution of monogamy in certain rhodent species.

Vasopressin and the Hypothalamic-Pituitary Axis

The hypothalamic paraventricular nucleus contains two categories of neurosecretory neurons capable of synthesizing vasopressin. We have already referenced the magnocellular neurons that project their axons through the internal zone of the median eminence to the neurohypophysis. The following refers to a subpopulation of parvocellular paraventricular neurons capable of co-synthesizing CRH and vasopressin.

Vasopressin in Corticotropin Releasing Hormone Neurons

Under resting conditions, most paraventricular nucleus neurons that synthesize CRH neurons do not normally contain detectable levels of vasopressin. The expression of their vasopressin genes can be rapidly switched on by a variety of acute, repeated, or chronic challenges. Information revealing the presence of such challenges or stressors may reach the parvocellular neurons via several neuronal pathways: (1) descending inputs from circumventricular organs in the lamina terminalis along the anterior wall of the third cerebral ventricle, signaling information about osmolality, cytokine, and hormonal molecules in the circulation; (2) projections from diverse limbic system areas, conveying inputs related to psychological or emotional stress; (3) ascending inputs from catecholaminergic

neurons in the dorsomedial and ventrolateral medulla, relaying information related to autonomic and cardiovascular regulation; and (4) unspecified local intrahypothalamic innervations, including recurrent feedback pathways among these neurons themselves. Presumably the content and meaning of the information conveyed via different pathways reflects stimulus specificity and will engage different populations of paraventricular neurons to respond, thereby accounting for the different patterns of vasopressin mRNA and hnRNA expressed among populations of hypothalamic neurons in response to different stress paradigms.

The signaling that regulates the expression of the vasopressin gene in these cells remains to be defined. Distinct from their magnocellular counterpart, these parvocellular neurons contain type 2 glucocorticoid receptors that mediate a direct inhibitory effect of corticosteroids on neuropeptide gene expression. The removal of the adrenals, the source of this circulating corticosterone, triggers vasopressin synthesis and upregulates CRH synthesis in populations of parvocellular paraventricular nucleus neurons; the peptides are co-packaged into dense core vesicles, transported to the zona externa of the median eminence, released into the hypophysial portal circulation, and carried (as hormones) into the anterior pituitary.

Vasopressin in the Anterior Pituitary

In the anterior pituitary, vasopressin potentiates the CRH-induced release of ACTH by binding to specific V_{1b} receptors on the corticotrophs, activating the inositol phospholipid cycle and protein kinase C, mobilizing calcium from intracellular stores, and triggering the release of ACTH from the rapid turnover pool. Genetically vasopressin-deficient Brattleboro rats have a blunted corticosterone response to certain forms of stressful stimuli, presumably reflecting the contribution normally made by vasopressin to the ACTH released by a stressor.

See Also the Following Articles

Corticotropin Releasing Factor (CRF); Neuroendocrine Systems; Oxytocin; Paraventricular Nucleus; Circadian Clock Genes as Modulators of Sensitivity to Genotoxic Stress.

Further Reading

Antoni, F. A. (1993). Vasopressinergic control of pituitary adrenocorticotropin secretion comes of age. *Frontiers in Neuroendocrinology* **14**, 76–122.

Barberis, C., Mouillac, B. and Durroux, T. (1998). Structural bases of vasopressin/oxytocin receptor function. *Journal of Endocrinology* **156**, 223–229.

Cole, R. L. and Sawchenko, P. E. (2002). Neurotransmitter regulation of cellular activation and neuropeptide gene expression in the paraventricular nucleus of the hypothalamus. *Journal of Neuroscience* **22**, 959–969.

Kalsbeek, A. and Buijs, R. M. (2002). Output pathways of the mammalian suprachiasmatic nucleus: coding circadian time by transmitter selection and specific targeting. *Cell Tissue Research* **309**, 109–118.

Landry, M., Vial-Porcile, E., Hökfelt, T., et al. (2003). Differential routing of coexisting neuropeptides in vasopressin neurons. *European Journal of Neuroscience* **17**, 579–589.

Ludwig, M. and Pittman, Q. J. (2003). Talking back: dendritic neurotransmitter release. *Trends in Neuroscience* **26**, 255–261.

Péqueux, C., Keegan, B. P., Hagelstein, M-T., et al. (2004). Oxytocin- and vasopressin-induced growth of human small-cell lung cancer is mediated by the mitogen-activated protein kinase pathway. *Endocrine-Related Cancer* **11**, 871–885.

Raggenbass, M. (2001). Vasopressin- and oxytocin-induced activity in the central nervous system: electrophysiological studies using in-vitro systems. *Progress in Neurobiology* **64**, 307–326.

Sabatier, N., Richard, P. and Dayanithi, G. (1998). Activation of multiple intracellular transduction signals by vasopressin in vasopressin-sensitive neurones of the rat supraoptic nucleus. *Journal of Physiology* (London) **513**(3), 699–710.

Urban, I. J. A., Burbach, J. P. H. and DeWeid, D. (1998). *Progress in brain research. Vol. 119: Advances in brain vasopressin.* Amsterdam: Elsevier.

Young, L. J. and Wang, Z. (2004). The neurobiology of pair bonding. *Nature Neuroscience* **7**, 1048–1054.

Zingg, H. H., Bourque, C. W. and Bichet, D. G. (1998). *Advances in experimental medicine and biology. Vol. 449: Vasopressin and oxytocin: molecular, cellular and clinical advances.* New York: Plenum Press.

Vietnam Veterans, Postwar Experiences and Health Outcomes

J A Boscarino
Mount Sinai School of Medicine, New York, NY, and Geisinger Health System, Danville, PA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J A Boscarino, volume 3, pp 656–661, © 2000, Elsevier Inc.

The Vietnam War
Vietnam Veteran Studies
Psychobiological and Biomedical Findings
Studies of Other Veteran Populations after Vietnam
Implications for Medical Sciences, Research, and Treatment
Conclusion

Glossary

Autoimmune disease	A disease produced when the body's normal tolerance of its own antigenic markers on the surface of cells is lost.
Cohort study	An observational study design commonly used in epidemiological studies of populations exposed to pathogens that could result in future diseases.
Diagnostic Interview Schedule (DIS)	A mental health diagnostic interview based on standard psychiatric nomenclature that made population mental health surveys and population-level estimates possible.
Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III)	The third edition of standard psychiatric nomenclature used to classify all emotional and mental health disorders in 1980 and the edition in which the definition of PTSD was included.
Heritability	The proportion of variance of the liability to a disease that is due to shared genetic factors.
Hypothalamic-pituitary-adrenal (HPA) system	Three key anatomic structures that collectively participate in hormonal responses to stressors by regulating human neuroendocrine functions and, hence, the body's adaptation to environmental stressors.
National Vietnam Veterans Readjustment Study (NVVRS)	A national cross-sectional survey of male and female Vietnam veterans and nonveterans conducted to determine the prevalence and etiology of psychosocial readjustment problems and mental disorders among Vietnam veterans.
Operation Ranch Hand	The U.S. military's operational code name for the application of herbicides (chiefly agents orange, blue, green, pink, purple, and white) in Vietnam by aircraft in order to destroy enemy cover and food crops.
Operation Ranch Hand Study	A follow-up study of the health status of the men involved in the Operation Ranch Hand herbicide operations.
Posttraumatic stress disorder (PTSD)	A specific disabling psychiatric disorder that develops after exposure to psychologically traumatic events; the disorder could last for decades. Symptoms include the reexperiencing of the event, avoidance of stimuli related to the event, and persistent symptoms of increased arousal associated with the event.
Psychoneuroendocrinology	A field of medical research that focuses on the association between neuroendocrine function and psychological status, especially as this relates to steroidogenic biological alterations and psychiatric illness.
Syndromes	A group of symptoms of a disordered function related to one another by means of a common pathology.
Vietnam Experience Study (VES)	A national cohort study of male, U.S. Army Vietnam-theater and non-Vietnam-theater veterans, conducted to determine the prevalence and etiology of diseases and medical conditions associated with Vietnam service.
Vietnam Era Twin (VET) Registry	A registry composed of male twin pairs who both served in the U.S. military between 1965–1975 in order to study the effect of genetic factors on health status.

The Vietnam War

Although U.S. involvement in the Vietnam War officially ended over 30 years ago for the 3.14 million men and 7200 women who served there, the impact of their postwar experiences altered society in many ways. The postwar experiences of Vietnam veterans were different from veterans of many previous wars. For example, unlike earlier U.S. wars in the twentieth century, the Vietnam veteran returned home from a conflict that was controversial and unpopular. Furthermore, many veterans were draftees who often came from lower socioeconomic (SES) households. In addition, of all men eligible for the military draft, only a specific subgroup was primarily inducted, due to the Selective Service System's deferment policy, which generally exempted men from higher SES groups. Within this context, it became clear to at least some health-care professionals during the war

that many returning veterans were experiencing postwar adjustment problems. During this period, the work of Charles Figley stood out as a classic in stress research because it documented the emotional and psychological trauma many veterans began to manifest by war's end. Based on Figley's work, and that of others, health-care researchers and medical professionals in the late 1970s formulated specific etiological models of the pathogenesis of mental disorders among Vietnam veterans. They also contributed to the development of case definitions to characterize these syndromes. These efforts eventually culminated in the inclusion of posttraumatic stress disorder (PTSD) in the *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn.; DSM-III). Once this was achieved, more accurate assessments of the scope and nature of traumatic stress (and related disorders) among Vietnam veterans and others were possible. Finally, studies of the postwar experiences of Vietnam veterans have led to many research developments and have accelerated the accumulation of knowledge related to the psychology and biology of PTSD, including the possible long-term medical and health consequences of this syndrome.

Although clinical documentation of the adverse psychological effects of combat exposure goes back at least as far the U.S. Civil War, the etiology of these syndromes were neither clearly understood nor well defined. Ninety years ago during World War I, soldiers with acute adverse psychological reactions after combat were believed afflicted with shell shock, thought to result from the concussion effects of artillery bombardments. During World War II, combat fatigue was used to categorize soldiers exhibiting acute adverse psychological reactions after combat because it was thought that exhaustion played a role in this condition. From the 1950s to 1970s, the term war neurosis often was used to broadly characterize soldiers' adverse psychological reactions after combat. For example, the psychiatric literature during this period associated a wide range of postwar psychological problems with combat exposure, including irritability, jumpiness, disturbed sleep, hysteria, disorientation, and panic attacks. However, for years the long-term effects of combat exposure remained elusive, in part because of a poor case definition and in part because reliable and valid observational studies were not available until years after World War II. After reviewing these and other findings related to natural disaster and concentration camp survivors, Dohrenwend concluded that there was clear and compelling evidence that many individuals develop functional psychiatric disorders after extreme stress exposures that had not been present before these experiences, a discovery that had been made earlier by Freud.

Vietnam Veteran Studies

Initial Research

Initial studies of the postwar experiences of Vietnam veterans provided the first clinical and scientific evidence linking combat exposures in Vietnam to postdischarge problems and adjustment difficulties also observed among previous veterans. However, in part because Vietnam veterans often came from lower SES groups, the postwar status of these veterans was controversial. In addition, many early studies were methodologically flawed. One difficulty was that these studies were limited because of biases associated with using nonrepresentative samples. Another problem was the use of nonstandardized mental health measures or the use of measures that assessed only mental health symptoms. Thus, although clinicians had observed that many individuals exposed to combat exhibited certain postwar syndromes afterward (including hyperalertness, exaggerated startle responses, sleep disturbances, and other symptoms) linking these to combat exposure was problematic because of existing methodological shortcomings. However, when DSM-III was being developed in the late 1970s, clinicians and others involved with Vietnam veterans were successful in eventually incorporating these syndromes into DSM-III under the diagnostic nomenclature defining PTSD, originally labeled the post-Vietnam syndrome and later conceptually broadened to include other types of trauma, such as sexual abuse. Once PTSD was included in DSM-III, undertaking large-scale, standardized surveys of Vietnam veterans was feasible and had other research benefits. During this early research period, Robert Laufer's Legacy Study, although not based on a true probability study, stood out as a landmark investigation. This is because it included over 1000 subjects (both Vietnam veterans and civilians) and used several standard symptom scales common in the 1960s and 1970s. Thus, at the time, this study provided the best assessment of the mental health status of Vietnam veterans and served as a benchmark for future studies.

Later Research

As part of Public Law 98–160, the U.S. Congress in 1983 mandated that research on Vietnam veterans be undertaken to determine "the prevalence and incidence of post-traumatic stress disorder and other psychological problems in readjusting to civilian life." In addition to Public Law 98–160, the U.S. Congress passed other public laws mandating studies of the health effects of Vietnam service. As a consequence of these new laws, several well-designed

cohort studies of Vietnam veterans were undertaken during the 1980s. These studies avoided many of the shortcomings of previous research, chiefly because of advancements in sampling and measurement that evolved from earlier research. One of these included the availability of the Diagnostic Interview Schedule (DIS), which for the first time permitted the gathering of DSM-III psychiatric diagnoses by means of population surveys, something previously not possible. This later research confirmed that Vietnam combat veterans had higher rates of postwar adjustment difficulties, mental health disorders, medical morbidity, and postwar mortality than noncombat veterans or comparable nonveterans. Most important, however, these studies indicated that the postwar adjustment difficulties and health problems experienced by these veterans were often due to combat exposures in Vietnam, not to the selection biases or measurement inadequacies that had affected earlier studies. Among the later studies, the National Vietnam Veterans Readjustment Study (NVVRS), which involved over 2000 Vietnam-theater and nontheater veterans (in addition to hundreds of civilian nonveterans), is considered one of the most comprehensive psychosocial assessment to date. The NVVRS indicated that 15% of male Vietnam veterans were current PTSD cases (9% of female veterans) and that 31% of male veterans (27% of female veterans) had PTSD during their lifetimes. Those with PTSD also were more likely to have other stress-related psychiatric disorder, such as depression and anxiety disorders, as well as many other postwar adjustment problems. The NVVRS further indicated that the PTSD-positive veterans often had profoundly disrupted lives in almost every domain of life, including in employment and family relationships. Furthermore, the prevalence of PTSD and other postwar psychological problems were found to be significantly higher among those with greater combat exposure and among Hispanic and African American veterans.

However, as significant as the psychosocial consequences of the war was for Vietnam veterans, their postwar experiences went beyond psychosocial outcomes. For example, the Vietnam Experience Study (VES) indicated that Vietnam veterans had higher rates of postwar mortality in the first 5 years after discharge, primarily due to suicides, homicides, drug overdoses, and motor vehicle accidents. Furthermore, VES findings related to the postwar medical morbidity experienced by these veterans confirmed that Vietnam-theater veterans as a group had higher rates of health-care use and reported themselves to be in poorer health than veterans without Vietnam service. In addition, theater veterans also tended to have higher rates of hearing loss, evidence of past hepatitis B infection, and higher thyroid-stimulating hormone levels after discharge. In addition, when the postwar health status of Vietnam veterans was examined in relation to whether the veteran had PTSD, PTSD-positive veterans had substantially higher (i.e., 50–150% greater) postwar rates of many major chronic diseases, including circulatory, nervous system, digestive, musculoskeletal, and respiratory diseases, even after controlling for the major risk factors for these conditions. For example, 25% of PTSD-positive veterans reported physician-diagnosed circulatory diseases nearly 20 years after the service (vs. 13% of PTSD-negative veterans) and 19% reported physician-diagnosed nervous system disorders (vs. 6%). Altogether, 68% of PTSD-positive Vietnam veterans reported the occurrence of a chronic disease-related medical condition 20 years after Vietnam service (vs. 48% of PTSD-negative Vietnam veterans). Finally, PTSD-positive veterans also were significantly more likely to show electrocardiographic evidence of myocardial infarction and to have abnormally high white blood cell counts ($>11{,}000/mm^3$) and other immune system abnormalities 20 years after military service.

The most compelling evidence, however, linking PTSD to adverse health outcomes was a VES study that examined the causes of death among 15,288 male U.S. Army veterans 30 years after their military service (see **Table 1**). The analyses adjusted for race, Army volunteer status, Army entry age, Army discharge status, Army illicit-drug abuse, intelligence, age, and for cancer mortality, also pack-years of cigarette smoking. The findings indicated that adjusted postwar mortality for all, cardiovascular, cancer, and external causes of death (e.g., motor vehicle accidents, accidental poisonings, suicides, homicides, and injuries of undetermined intent) was associated with PTSD among Vietnam-theater veterans ($N = 7924$), with hazards ratios (HRs) of 2.2 ($p < 0.001$), 1.7 ($p = 0.034$), 1.9 ($p = 0.018$), and 2.3 ($p = 0.001$), respectively. For Vietnam-era veterans with no Vietnam service ($N = 7364$), PTSD was associated with all causes of mortality (HR = 2.0; $p = 0.001$). PTSD-positive Vietnam-era veterans also appeared to have an increase in external-cause mortality as well (HR = 2.2; $p = 0.073$). This study essentially suggested that Vietnam veterans with PTSD were at approximately twice the risk of postwar death from multiple causes. Other postwar studies were also conducted, such as the Operation Ranch Hand Study, which sought to assess the postwar health effects of exposure to herbicides used in Vietnam. Others include the Vietnam Era Twin (VET) studies, which focused on the genetic liabilities for PTSD, substance abuse, and other mental health disorders, not just the impact of Vietnam service.

Table 1 Cox proportional hazards regressions: crude and adjusted hazard ratios by veteran and PTSD status[a]

Veteran status	*All-cause mortality (total deaths = 820)*			*Cardiovascular mortality (total deaths = 241)*			*Cancer mortality (total deaths = 188)*			*External-cause mortality (total deaths = 175)*		
	HR	*95% CI*	*P value*	*HR*	*95% CI*	*P value*	*HR*	*95% CI*	*P value*	*HR*	*95% CI*	*P value*
All veterans[b]												
PTSD (unadjusted)	2.5	2.1–3.0	<0.001	1.7	1.1–2.6	0.010	1.8	1.1–2.8	0.013	2.7	1.8–4.0	<0.001
PTSD (adjusted)[c]	2.1	1.7–2.6	<0.001	1.6	1.1–2.4	0.027	1.5	1.0–2.4	0.075	2.3	1.5–3.5	0.001
Vietnam-era veterans[d]												
PTSD (unadjusted)	2.6	1.7–3.8	<0.001	1.3	0.5–3.6	0.57	1.1	0.4–3.6	0.84	2.9	1.3–6.7	0.012
PTSD (adjusted)[c]	2.0	1.3–3.0	0.001	1.2	0.4–3.4	0.69	0.9	0.3–3.1	0.92	2.2	0.9–5.2	0.073
Vietnam-theater veterans[e]												
PTSD (unadjusted)	2.5	2.0–3.2	<0.001	1.8	1.1–2.8	0.015	2.2	1.3–3.7	0.003	2.6	1.6–4.1	<0.001
PTSD (adjusted)[c]	2.2	1.7–2.7	<0.001	1.7	1.0–2.7	0.034	1.9	1.1–3.3	0.018	2.3	1.4–3.9	0.001

[a]HR = hazards ratio; CI = confidence interval; PTSD, = posttraumatic stress disorder. Adapted from J. A. Boscarino (2006). Posttraumatic stress disorder and mortality among US Army veterans 30 years after military service. *Annals of Epidemiology* **16**, 248–256. Reproduced with permission by Elsevier Publishing.
[b]*N* = 15,288; person-risk years = 229,565; total PTSD cases = 1050.
[c]All models adjusted for race, Army volunteer status, Army entry age, Army discharge status, Army illicit drug use, age at interview, and intelligence. For cancer mortality, models adjusted for the above variables, in addition to pack-years of cigarette smoking.
[d]*N* = 7364; person-risk years = 110,553; total PTSD cases = 214.
[e]*N* = 7924; person-risk years = 119,453; total PTSD cases = 836.

To date, however, the VES has provided the most compelling evidence related to the adverse health effect of combat service and PTSD during that conflict.

Psychobiological and Biomedical Findings

There are clinical reasons to expect alterations in neuroendocrine system functions in chronic PTSD cases because long-term changes in the HPA system and the sympathetic arm of the autonomic nervous system have been reported following exposure to extreme stress. Evidence suggests that the physiological arousal often observed during the recollection of traumatic events is associated alterations in the neuroendocrine functions linked to the sympathetic-adrenomedullary and HPA stress axes. In the case of PTSD, it is thought that these neuroendocrine alterations reflect the consequences of an extreme state of physiopsychological conditioning that occurs following severe stress exposures. Furthermore, although this conditioning response is initiated in the central nervous system (CNS), it is subsequently carried out by multiple endocrine mechanisms that have wide-ranging effects on the body and nervous system. Similar findings related to this physiological-emotional conditioning process have also been observed with animal models.

Several studies have reported lower cortisol levels among Vietnam veterans with PTSD, as well as increased catecholamine concentrations. Furthermore, one recent large-scale study, based on the VES and involving thousands of veterans, found not only that Vietnam-theater veterans with current PTSD had lower cortisol but that past combat exposure levels were associated with plasma cortisol concentrations in an inverse dose–response relationship (see **Figure 1**). In brief, research with Vietnam combat veterans suggests that PTSD-positive veterans tend to have lower cortisol levels but, paradoxically, higher catecholamine concentrations, together with heightened responses of the stress system to traumatic memories and other stimuli associated with the original trauma. It is known that chronicity and excessiveness of stress system activation can lead to pathogenesis, causing weight changes, depression, hypogonadism, and other systemic alterations, including increased metabolic load on the body's systems. Thus, the physiological findings for PTSD-positive Vietnam veterans suggest that past traumatic stress exposure has resulted in neuroendocrine system alterations that make veterans potentially susceptible to a host of chronic diseases. As noted, PTSD-positive Vietnam veterans were found to have higher postwar rates

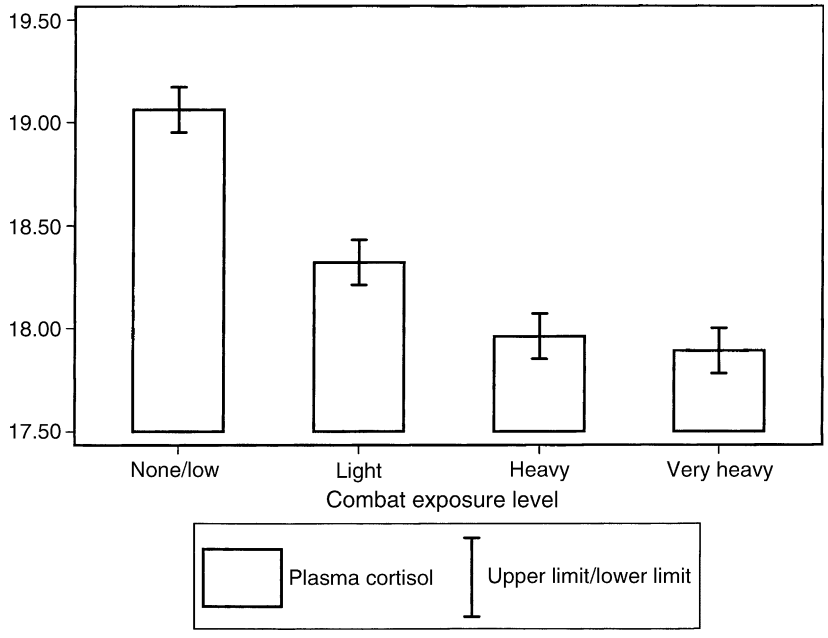


Figure 1 Adjusted plasma cortisol concentrations and standard errors bars (95% confidence limits) for Vietnam-theater veterans by combat exposure levels ($N = 2457$). Results shown are for morning, fasted plasma cortisol (in micrograms per deciliter). Combat level was measured by the Combat Exposure Index scale, a standard measure of combat exposure. Cortisol results are adjusted for demographic variables, alcohol use, drug use, smoking status, and body mass index using analysis of covariance. A trend test for cortisol by combat exposure level was statistically significant ($p = 0.0001$), confirming the inverse dose–response association between cortisol and combat exposure. Adapted from J. A. Boscarino (1996), Post-traumatic stress disorder, exposure to combat, and lower plasma cortisol among Vietnam veterans: findings and clinical implications, *Journal of Consulting and Clinical Psychology* **64**, 191–201. Reprinted with permission from the American Psychological Association.

of many chronic diseases, reinforcing this neuroendocrine-dysregulation disease hypothesis, which represents a cornerstone in human psychoneuroendocrinology research.

Related to HPA dysregulation, it has been suggested that PTSD-positive veterans may be at risk for autoimmune diseases as well. For example, although there have been inconsistencies, investigators, as previously noted, have found that individuals who developed PTSD, particularly men exposed to combat, appeared to have lower plasma cortisol concurrent with higher catecholamine levels. As suggested, Vietnam veterans with current PTSD not only had lower cortisol, but this had an inverse dose–response relationship with combat exposure. In addition, also as previously noted, research indicated that Vietnam veterans with current PTSD had clinically elevated leukocyte and T-cell counts, and similar findings have been reported in nonveteran studies as well. These veterans were also found to have other significant neuroendocrine abnormalities, such as elevated thyroid hormones. Consistent with these clinical findings, it has recently been reported that PTSD was in fact associated with several autoimmune diseases among Vietnam veterans, including rheumatoid arthritis and psoriasis. Again, analogous results have been reported in animal models as well. Of course, not all of the variance in disease outcomes should be associated directly with PTSD. Research with the VET Registry suggested that about one-third of the liability for PTSD was associated with genetic factors. Some of the same genes that make one liable for PTSD could also make one liable for specific diseases as well.

Studies of Other Veteran Populations after Vietnam

Following the first Persian Gulf War (PGW) in 1991, cohort studies were again used to evaluate the health impact of this conflict on war-theater military personnel. These studies are still in progress, and a number of them evolved from or were related to the original NVVRS and VES research designs. The PGW research to date has established a link between service in the PGW and mental illness, postservice injury, chronic fatigue syndrome, and other multisymptom conditions. The reasons for these associations are still being investigated, and the results from PGW research are being used to inform a new generation of veteran studies, related to personnel deployed in the Afghanistan and Iraq theaters of war following the September 11, 2001, attacks. For these recent conflicts, however, military personnel were being assessed both before deployment and after return, something rarely done in the past but often desired after service exposure had occurred. Finally, following the September 11 attacks in New York City (NYC) in 2001, a number of the investigative teams, funded by the National Institutes of Health (NIH) and the Centers for Disease Control (CDC) to assess the population impact of this event, used the prospective cohort method as well as measurement instruments that had been developed originally for Vietnam veteran population studies. Thus, the field of disaster epidemiology has emerged as a subspecialty within epidemiology, informed, at least in part, by the original work with Vietnam veterans over the previous decades.

Implications for Medical Sciences, Research, and Treatment

The postwar findings reported for Vietnam veterans documented that veterans exposed to combat have more stress-related psychosocial disorders, altered neuroendocrine functions, higher medical morbidity, and greater mortality years after the war. The scientific findings related to the postwar experience of the Vietnam veteran suggest that psychobiological models are probably required to understand the impact of this experience on the lives of those who served there. For example, it has been suggested that traumatic stressors act as unconditioned aversive stimuli that evoke severe physiological distress. Consequently, previously neutral external stimuli (e.g., the sound of a helicopter) and internal stimuli (e.g., physiological states) that accompanied the traumatic stressor, can function as conditioned stimuli capable of producing psychological and physiological distress when the traumatic stressors are no longer present. Experience with psychotherapy treatments suggests that when stress victims are gradually exposed to aversively conditioned stimuli the symptoms may decrease because desensitization to the stressor could occur. A psychotherapeutic technique often used in PTSD therapy with veterans is to teach the patient to verbalize the traumatic experience with the help of a support group consisting of other veterans. This process allows veteran to redefine the external and internal stimuli that affect psychological arousal, develop better self-control, and, with assistance, desensitize him- or herself to a range of thoughts, actions, and situations associated with the traumatic events. A more structured and currently widely used approach that appears to be effective is cognitive-behavioral therapy. Thus, psychosocial processes appear to be instrumental in treating a condition that is psychosocial in origin but that also has a physiobiological basis – that is, research with Vietnam veterans and

others tells us that PTSD has both a psychosocial and a physiobiological foundation.

The recognition of a multifactorial PTSD model is consistent with the report that both pharmacotherapy and cognitive-behavioral psychotherapy are effective in treating PTSD. In the case of pharmacotherapy, the pathophysiology of PTSD, in part, appears to involve the serotonergic and the noradrenergic systems; thus, drugs known to potentiate these mechanisms have been effective. In the case of cognitive-behavioral therapy, this approach has been found to be effective in reducing PTSD-related symptomotology by achieving desensitization to stressful stimuli, by increasing control of adversive arousals, by enhancing anxiety management, and by using other known behavioral-psychological mechanisms. Although the underlying interventional mechanisms differ for pharmacological versus cognitive-behavioral therapy (e.g., pharmacokinetic vs. psychological), the outcomes are similar – the psychopathology and underlying pathophysiology are reduced and fewer adverse patient symptoms are manifested, hence lowering the risk of other adverse outcomes. Thus, we expect that as trauma-related symptoms are reduced through treatment, the risk for other adverse outcomes will also decrease. Recent research has, in fact, confirmed this hypothesis in NYC following the September 11 attacks. For example, it has been reported that NYC adults who received emergency crisis counseling at work shortly after the attacks not only had better mental health outcomes but also had better outcomes in terms of binge drinking, alcohol dependence, and alcohol consumption, lowering the possibility of future adverse health outcomes.

Although the findings discussed here chiefly involve veterans, exposure to combat, they are applicable to other populations. Traumatic stress reactions occur among individuals exposed to extreme, life-threatening situations, including natural disasters, physical attacks, wartime combat, and other situations. In addition, it is likely that human exposure and response to stress take place along a stress continuum, with traumatic stress on one end of this continuum. Furthermore, genetic liability is also likely involved with respect to whom may succumb to PTSD. Finally, the postwar experiences of the Vietnam veteran suggest that more than psychosocial explanations are needed to understand PTSD and to treat it effectively. In addition to higher-order cognitive models, findings in Vietnam veteran suggest lower-order biological models are also necessary. Research also consistently suggests that those with less social support exhibit greater levels of traumatic stress, which may explain the high rates of PTSD among Vietnam veterans, given the low community support during and after the war. This finding has clinical implications for the future treatment of both combat and noncombat-related stress victims.

Conclusion

Psychoneuroendocriniology represents a growing area of medical research that focuses on the association between neuroendocrine function and psychological status, especially as this relates to steroidogenic biological alterations and psychiatric illness. Just as with other areas of medicine – orthopedics, plastic surgery, emergency medicine, and so forth – developments in

Table 2 Key biological correlates of PTSD among a national random sample of Vietnam veterans[a]

Biometric measure	*PTSD negative (%)*	*PTSD positive (%)*	*P value*
White blood cell count elevated	2.8	4.6	0.063
Total lymphocytes elevated	4.2	6.8	0.023
T lymphocytes elevated	2.9	6.2	0.001
B lymphocytes elevated	9.9	13.7	0.032
CD4 cells elevated	3.5	5.6	0.055
CD8 cells elevated	2.4	5.9	<0.001
Erythrocyte sedimentation rate elevated	6.1	9.6	0.013
Triiodothyronine (T_3) uptake elevated	7.9	10.8	0.06
Thyroxine (T_4) elevated	1.7	5.0	<0.001
Immunoglobulin-A elevated	16.4	23.5	0.001
Immunoglobulin-M elevated	1.2	3.4	0.001
Dehydroepiandosterone reduced	2.3	4.6	0.008
Cortisol/dehydroepiandosterone ratio elevated	4.4	7.1	0.023
N =	4139	323	–

[a]Data are from the Vietnam Experience Study and represent a national random sample of male U.S. Army veterans of the Vietnam War era, half of which served in Vietnam and half of which served elsewhere. PTSD cases represent both combat and noncombat cases. Assessment was done approximately 20 years after military service. PSTD, posttraumatic stress disorder.

military medicine and veterans' health have been translated into medical advances for other populations. For example, today, research related to the biological and medical consequences of anxiety disorders is more commonplace, as are epidemiological studies of other veteran populations and nonveteran studies involving exposure to human-made and natural disasters.

The medical promise of this body of research can be illustrated using the findings shown in Table 2. As shown, PTSD-positive veterans have a distinct clinical profile, suggesting an increased inflammatory response related to a broad range of potential pathophysiologies that could be associated with a host of diseases, including cardiovascular and autoimmune diseases. The reasons for the findings shown in Table 2 are not clear at this time; they may be related to biological, psychological, or behavioral factors (or a combination of these) associated with PTSD and research also suggests that genetic factors are indeed involved. What is clear is that this area of research holds promise for understanding the psychosomatic basis for the causes and prevention of many human diseases. In conclusion, despite of the suffering caused by Vietnam War, or perhaps because of it, the postwar experiences of the Vietnam veteran have resulted in significant advances in our knowledge about the nature and consequences of human stress exposures and will probably result in future medical advances that cannot currently be imagined.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Allostasis and Allostatic Load; Anxiety; Autoimmunity; Catecholamines; Combat Reaction, Chronic; Combat Stress Reaction; Combat, Acute Reactions to; Corticosteroids and Stress; Depression Models; Disease, Stress Induced; Gulf War Syndrome, Psychological and Chemical Stressors; Hypothalamic-Pituitary-Adrenal; Korean Conflict, Stress Effects of; Major Depressive Disorder; Nuclear Warfare, Threat of; Persian Gulf War, Stress Effects of; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of; Psychotherapy; Salivary Cortisol; War Stress in the Former Yugoslavia; War-Related Posttraumatic Stress Disorder, Treatment of; Depression, Immunological Aspects; Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder – Neurobiological basis for.

Further Reading

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd edn.). Washington, DC: American Psychiatric Association.

Archibald, H. C., Long, D. M., Miller, C., et al. (1962). Gross stress reaction in combat – a 15-year follow-up. *American Journal of Psychiatry* **119**, 317–322.

Boscarino, J. A. (1995). Post-traumatic stress and associated disorders among Vietnam veterans: the significance of combat and social support. *Journal of Traumatic Stress* **8**, 317–336.

Boscarino, J. A. (1996). Post-traumatic stress disorder, exposure to combat, and lower plasma cortisol among Vietnam veterans: findings and clinical implications. *Journal of Consulting and Clinical Psychology* **64**, 191–201.

Boscarino, J. A. (1997). Diseases among men 20 years after exposure to severe stress: implications for clinical research and medical care. *Psychosomatic Medicine* **59**, 605–614.

Boscarino, J. A. (2004). Association between posttraumatic stress disorder and physical illness: results and implications from clinical and epidemiologic studies. *Annals of the New York Academy of Sciences* **1032**, 141–153.

Boscarino J. A. (2006). Posttraumatic stress disorder and mortality among US Army veterans 30 years after military service. *Annals of Epidemiology* **16**, 248–256.

Boscarino, J. A., Adams, R. E., Foa, E. B., et al. (2006). A propensity score analysis of brief worksite crisis interventions after the World Trade Center disaster: implications for intervention and research. *Medical Care* **44**, 454–462.

Centers for Disease Control (1987). Postservice mortality among Vietnam veterans. *Journal of the American Medical Association* **257**, 790–795.

Centers for Disease Control (1988). Health status of Vietnam veterans. I: Psychosocial characteristics. *Journal of the American Medical Association* **259**, 2701–2707.

Centers for Disease Control (1988). Health status of Vietnam veterans. II: Physical health. *Journal of the American Medical Association* **259**, 2708–2714.

Chrousos, G. P. and Gold, P. W. (1992). The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. *Journal of the American Medical Association* **267**, 1244–1252.

Dohrenwend, B. P. (1975). Sociocultural and social-psychological factors in the genesis of mental disorders. *Journal of Health and Social Behavior* **16**, 365–392.

Figley, C. R. (1978). *Stress disorders among Vietnam veterans: theory, research and treatment.* New York: Brunner/Mazel.

Freud, S. (1953–1974). Introduction to psychoanalysis, war neuroses (1919). In: Strachey, J. (ed.) *Standard edition of the complete psychological works of Sigmund Freud*, (vol. 17), pp. 205–215. London: Hogarth Press and the Institute of Psychoanalysis.

Frey-Wouters, E. and Laufer, R. S. (1986). *Legacy of a war: The American soldier in Vietnam.* Armonk, NY: Sharpe.

Galea, S., Ahern, J., Resnick, H., et al. (2002). Psychological sequelae of the September 11 terrorist attacks in New York City. *New England Journal of Medicine* **346**, 982–987.

Hanson, F. R. (1949). The factor of fatigue in the neuroses of combat. In: Hanson, F. R. (ed.) *Special issue on combat psychiatry. Army Medical Bulletin* **9**, 147–150.
Helzer, J. E., Robins, L. N. and McEvoy, L. (1987). Posttraumatic stress disorder in the general population: findings of the Epidemiologic Catchment Area Study. *New England Journal of Medicine* **317**, 1630–1634.
Hoge, C. W., Auchterlonie, J. L. and Milliken, C. S. (2006). Mental health problem, use of mental health services, and attrition from military service after returning from deployment to Iraq or Afghanistan. *Journal of the American Medical Association* **295**, 1023–1032.
Hyams, K. C., Wignall, F. S. and Roswell, R. (1996). War syndromes and their evaluation: from the U.S. Civil War to the Persian Gulf War. *Annals of Internal Medicine* **125**, 398–405.
Institute of Medicine (2001). *Veterans and agent orange, update 2000*. Washington, DC: National Academy Press.
Keane, T. M., Zimering, R. T. and Caddell, J. M. (1985). A behavioral formulation of posttraumatic stress disorder in Vietnam veterans. *Behavior Therapist* **8**, 9–12.
Kulka, R. A., Schlenger, W. E., Fairbank, J. A., et al. (1990). *Trauma and the Vietnam War generation: report of findings from the National Vietnam Veterans Readjustment Study*. New York: Brunner/Mazel.
Mason, J. M., Giller, E. L., Kosten, T. R., et al. (1988). Elevation of urinary norepinephrine/cortisol ratio in posttraumatic stress disorder. *Journal of Nervous and Mental Diseases* **176**, 498–502.
McEwen, B. S. (2000). Allostasis and allostatic load: implication for neuropsychopharmacology. *Neuropsychopharmacology* **22**, 108–124.
Polner, M. (1971). *No victory parades: the return of the Vietnam veteran*. New York: Holt, Rinehart, Winston.
Robins, L. N., Helzer, J. E., Croughan, J., et al. (1981). National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics, and validity. *Archives of General Psychiatry* **38**, 381–389.
Salmon, T. W. (1919). War neuroses and their lesson. *New York Medical Journal* **109**, 993–994.
True, W. R., Rice, J., Eisen, S. A., et al. (1993). A twin study for genetic and environmental contributions to liability for posttraumatic stress sympotoms. *Archives of General Psychiatry* **50**, 257–264.
van der Kolk, B. A. and Saporta, J. (1993). Biological response to psychic trauma. In: Wilson, J. P. & Raphael, B. (eds.) *International Handbook of traumatic stress syndromes*, pp. 25–33. New York: Plenum Press.
Wolkowitz, O. M. and Rothschild, A. J. (eds.) (2003). *Psychoneuroendocrinology*. Washington, DC: American Psychiatric Publishing.
Yehuda, R. (2002). Current status of cortisol findings in post-traumatic stress disorder. *Psychiatric Clinics of North America* **25**, 341–368.

Vietnam War, Stress Effects of *See:* Captivity, Adaptation to; Captivity, Recovery from; Combat Reaction, Chronic; Combat, Acute Reactions to; Coping and Stress: A Lens and Filter Model; Vietnam Veterans, Postwar Experiences and Health Outcomes; Defense Services, Stress in.

Violence

E K Englander
Bridgewater State College, Bridgewater, MA, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by E K Englander, volume 3, pp 662–668, © 2000, Elsevier Inc.

Types of Biological Influences
Summary

Glossary

Adoption studies	Research wherein a child with one set of biological parents and a different set of psychological parents is studied.
Biological environmental influences	Events that affect a person biologically but are not encoded into the person's DNA.
Concordance rates	The frequency with which one twin is diagnosed with an illness, when his or her twin is also diagnosed with the same illness.
Cortisol	

	The hormone that regulates the body's reaction to stress.
Genetic influences	The biological blueprints for behavior that are contained in a person's chromosomes.
Imitation	A theory that states that children may learn to be violent by imitating an adult's aggression.
Minimal brain dysfunction or minimal brain damage	Very low levels of brain dysfunction, theoretically affecting behavior, emotions, learning, or memory.
Neurotransmitters	Chemicals secreted by the brain, some of which have been linked to aggression.
Paranoid misperceptions	Attitudes or belief systems that misinterpret ambiguous circumstances as signs of danger when none actually exists.
Resilient state	The state in which a person is likely to resist developing a disorder.
Serotonin	One of the monoamines, low levels of which have been implicated as related to aggression.
Testosterone	One of the androgens. A male sex hormone that is associated with aggression.
Vulnerable state	The state in which a person is likely to develop a disorder.

Violent crimes can be as baffling as they are repelling. In England, two 10-year-old boys lured a 2-year-old away from his mother at a shopping mall; they took him down to the local railroad tracks, where they beat him to death and left his body to be run over by a train. In the United States, a group of teenagers shot and killed a man whose car they were stealing – even though the man had willingly handed over his car keys and was not resisting the theft in any way. In Massachusetts, a group of young teenage schoolgirls planned to murder their English teacher because she was too strict.

To understand the causes of violence, it is not always enough to understand the external circumstances that often drive other types of criminal behavior. It is true that violence sometimes results from factors such as an individual's desire for financial gain or an individual's repeated exposure to violent behavior in his or her social environments. Sometimes, however, violence also happens in an apparently motiveless fashion. Violence is never truly without motive, but the motives may be so complex and elusive that the violence appears motiveless. In all cases, but particularly in cases of violence that appear to have no motive, internal or individual factors may be critical for understanding the cause of such behavior. A variety of different biological and psychological influences and mechanisms have been considered over the years. This article summarizes them and attempts to construct a more comprehensive model of the causes of violent behavior.

Types of Biological Influences

There exists a mistaken tendency to use the terms biological and genetic interchangeably. In fact, genetic influences are only one type of biological influences on behavior. There are (at least) two different types of biological influences: genetic influences and biological environmental influences. Genetic influences refer to the blueprints for behavior that are contained in a person's chromosomes. Chromosomes contain deoxyribonucleic acid (DNA), the genetic material a person inherits from his or her biological parents, which is referred to as their genotype. Biological environmental influences, unlike genetic influences, are events that affect a person biologically but that are not encoded into the person's DNA. For example, consider a head injury from a car accident, which subsequently changes the victim's personality. The head injury is unmistakably a biological environmental influence. No genotype determines that someone will have a car accident; however, such an accident, and its accompanying head injury, can still have an important biological effect on the person's behavior.

Although learning is clearly related to violent behavior, learning theories alone cannot fully explain violence in human beings. As an example, consider the principle of imitation. It states that children may learn to be violent by imitating an adult's aggression. According to this principle, one would expect that children who grow up watching violent parents would be violent themselves. As the theory predicts, a higher proportion of children who witness violent parents are violent, in comparison to other children; however, despite their higher risk, many children of violent parents do not, in fact, behave aggressively as adults. Some children who are exposed to violent psychosocial environments do become violent; many other children who are similarly exposed do not. This fact has led researchers to coin the terms resilient and invincible to describe children who survive and cope well despite terrible circumstances.

What makes one child resilient and another child vulnerable? Researchers have identified a variety of differences, including biological differences, that distinguish resilient and vulnerable children. It is important to note that such biological differences do not entirely explain why one child is vulnerable and another is resilient; rather, biology appears to be an important part of the determination. It is reasonable to assume that different biological influences will play

different roles in the determination of vulnerability and resiliency.

One important area that researchers have studied focuses on minimal brain damage. Also called minimal brain dysfunction, minimal brain damage does not look like major brain damage. A person with minimal brain dysfunction will not have obvious brain damage; however, minimal dysfunction may affect behavior, emotions, learning, or memory – that is, the more subtle (or higher) functions of the brain. For example, although major brain damage might result in severe mental retardation, minor brain dysfunction might result in a learning disability that would not necessarily affect intelligence in any way (e.g., dyslexia). It has been repeatedly demonstrated that minimal brain dysfunction is related to problems such as learning disorders and hyperactivity. Those relationships emerged as a rationale for studying a link between minimal brain dysfunction and violence, but how to study such a link is somewhat controversial in the field. Some researchers have examined violent individuals to see if they have high levels of minor physical anomalies, while others have given aggressive and violent men neuropsychological examinations. Still other psychologists have combed through the medical histories of violent individuals, seeking unusually high levels of brain injury in their histories. Although different researchers have taken different tacks, their findings have been consistent. However one chooses to measure minimal brain dysfunction, a host of research suggests that it is indeed related to persistent aggression and violence. Further, although both major and minor brain damage may cause violence, the weight of the evidence suggests that the most important type of brain dysfunction in people who are chronically violent and criminal is minimal brain dysfunction. Although it is almost certain that most people with minimal brain dysfunction will never be violent, a noticeably high number of repeat violent offenders show some level of brain dysfunction (usually minimal).

Neurotransmitters

There are many different types of neurotransmitters, and they tend to specialize in function. How much, or how little, of any particular neurotransmitter chemical is in your brain is an important determinant of many types of behavior. Both biology and psychosocial environment affect the levels of different neurotransmitters, and thus affect behavior. A logical question, therefore, is whether violent behavior is related (at least in part) to the level of certain neurotransmitters in the brains of violent people.

Monoamine neurotransmitters have been most significantly linked to aggression. Of the monoamines, low levels of serotonin have been implicated most strongly in recent research. Several studies have found consistent relationships between reduced serotonin activity and aggression and hostility in both boys and men. As is often the case in such research, the relationship was strongest among males, regardless of their age. This set of studies is interesting because it establishes such relationships among psychiatrically well patients as well as among behaviorally disordered individuals.

Hormones

Two hormones have been the focus of most research on aggression and criminality: testosterone and cortisol. Testosterone is one of the male sex hormones, called androgens. Both males and females secrete androgens, but males secrete a much greater quantity than females. Androgens are frequently cited as an important cause of aggressive behavior, particularly intermale aggression. This area of study is not completely consistent, however, and while some research points to aggression and dominance/hostility as being related to testosterone, other research fails to find that dominance/hostility measures are related to sex hormones. Some research notes a testosterone–aggression relationship in both males and females; other research finds it only for males. To muddy the waters even further, researchers who follow those experiments involving reduced testosterone functioning do not note that this is a clearly effective means of stopping aggression. In addition, psychosocial researchers suggest that high testosterone might result in kids being treated differently socially, since high testosterone is associated with being bigger and stronger. Androgens, therefore, have been implicated, but not strongly, in the cause of aggression.

A second hormone that has been implicated more recently is cortisol. Cortisol is the hormone that regulates the body's reaction to stress. It is involved with the immune system and with sex hormones as well. A few studies have linked low levels of cortisol with a tendency to be aggressive. Although this should be characterized as an exploratory area of research, it is interesting to note that low cortisol has even been found in females with conduct disorders.

Genetics

The majority of researchers agree that violence does not appear to be a behavioral tendency that is transmitted via a simple, directly acting gene. In any case, it is clear that if genetic bases for aggression and violence exist, they are clearly mutable and changeable by a person's psychological environment; the old idea that genetics determines one's behavior absolutely is clearly mistaken.

The case for genetics as a cause of aggression and adult criminal violence has to be pieced together from a number of different areas of study. First, we know that children with violent parents do have a higher tendency to be violent themselves. Second, data from twin studies show the same trend.

Twin studies are natural genetic laboratories; they are studies that typically compare identical (monozygotic) to fraternal (dizygotic) twins. Identical twins are the only human beings alive who are genetically identical. Some psychological disorders have much higher concordance rates for monozygotic twins. For example, schizophrenia has a concordance rate of 50%; if one identical twin has schizophrenia, then there is a 50% chance that the other twin will also have schizophrenia. Similar, and even higher, rates of concordance for aggressive behavior have been noted among twin pairs.

Despite these family studies finding a reasonably strong genetic link for aggression, some clarification is important. Specifically, the field needs to make the critical distinction between biological heredity versus family psychology. Parents provide a psychological, as well as a biological, environment for their children. Likewise, twin studies fail to tease apart the effects of psychosocial environment and genetics. In addition to sharing genes, twins also often share an environment that treats them similarly.

For this reason, adoption studies (in which a child has one set of biological parents and a different set of psychological parents) may do a better job of teasing apart heredity and psychological environment. Earlier classic studies, which examined children prospectively in long-term studies, found that children seemed to follow their adoptive parents' behavior, rather than their genetic parents (suggesting a stronger link with psychological environment than genetics). More recent research has also found that biological siblings had a higher concordance for aggression, in comparison to adoptive siblings.

Both family studies and adoption studies suggest a definite degree of heritability for aggression, but there is a third body of research that is relevant here: studies that examine the heritability of childhood precursors of violence in adulthood. This body of research examines childhood disorders that are highly related to violence and studies whether these childhood disorders are heritable. For example, attention-deficit/hyperactivity disorder and conduct disorders are both childhood disorders with strong ties to adult violence, and both disorders appear to be at least partially heritable. The heritability of a disorder such as violence may therefore proceed through the pathways of other, co-existing disorders, or, alternatively, some childhood disorders may be precursors to adult violence in a child who has inherited that behavioral tendency.

In summary, much of the data on the genetic basis of violence is suggestive of a significant, but not total or direct, heritability.

Hostile Interpretation of Ambiguous Events, or Cognitive Distortions

In 1996, a 17-year-old boy attacked a total stranger, unprovoked, with an ax. He knocked on the man's door, and when it was opened he attempted to attack before he was driven off by his intended victim. His defense in court? That he was mentally deranged and unusually paranoid due to late-stage Lyme disease. Before he became ill, this boy apparently had no behavioral problems, but his Lyme disease caused him to develop the unusual, but not unheard of, symptom of paranoid delusional misperceptions.

Cases this like are certainly rare. But are paranoid misperceptions rare among violent individuals? One theory suggests that such biased misperceptions may in fact help explain why some people are violent. Indeed, as the theory goes, under certain circumstances, aggression may be a normal and adaptive response. This may particularly be the case when an individual is faced with hostile or threatening forces, and violence is perceived as a necessary response. Given that a hostile environment might provoke an aggressive response in almost anyone, some researchers have investigated the possibility that violent individuals are not people who inappropriately choose violence; rather, they are people who inappropriately perceive hostility in situations where most people would not. That is, perhaps it is not the response that is skewed, but the perception of the environment.

How do aggressive children perceive the environment, and are their perceptions uniform? Studies find that aggressive children have two perceptual tendencies that nonaggressive children lack. Aggressive children are more likely to believe that other people had hostile intentions, and they are more likely to evaluate the results of aggression positively. Classic research has identified this thought pattern and has found that hostile biases were particularly characteristic of reactive-aggressive children. Researchers have similarly found that aggressive boys are significantly better at recognizing aggressively slanted comments. Studies have noted that children with memories that emphasize hostility process information in a more negatively biased way and are more likely to develop stable negative cognitive biases and stably aggressive behavior.

This line of study sheds some light on why it may be so difficult to treat violent offenders. Trying to teach such offenders to be nonviolent may in fact be trying

to teach them to react nonviolently to what they perceive to be intensely threatening situations. For example, imagine that you are standing in a field holding a gun, and an elephant is charging toward you in a rage at full speed with his tusks aimed right at you. Trying to teach violent offenders to be nonviolent might be like trying to teach them not to shoot that elephant, even though there is every indication that that elephant is deadly.

Other research has examined different types of violent offenders and has found cognitive misinterpretations and biases that appear to be specific to the type of offense. For example, rather than making general, nonspecific hostile misinterpretations, adults who commit child abuse tend to misinterpret children's behavior. Specifically, they tend to regard children's misbehaviors as more intentional than they really are – possibly paving the way to righteous anger and abuse. Sexually aggressive men, rather than suffering from blanket distortions, may differ from other men primarily in their distortion of women's communications. It seems clear that a wide variety of cognitive misperceptions seems to pave the way for a variety of different types of aggressive and violent behavior, and may help explain the difficulty of effectively treating such behavior.

Summary

Strong evidence exists to support the theory that individual differences can affect a person's risk for developing chronically violent behavior. These risks appear to interact with psychosocially learned factors and may differ significantly from individual to individual. The mechanism by which these associated risk factors contribute to violence is less well established. Research has identified heritability, hormones, minimal brain dysfunction, and biased cognitive misperceptions as likely areas of dysfunction related to violence and aggression. Future research will focus on effective prevention efforts and a better understanding of how risk factors influence individuals' behaviors.

Acknowledgments

This article was adapted, with permission, from *Understanding Violence* (Lawrence Erlbaum & Associates, Inc., 2003).

See Also the Following Articles

Aggression; Aggressive Behavior; Cortisol Awakening Response.

Further Reading

Crick, N. and Dodge, K. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development* **67**(3), 993–1003.

Englander, E. K. (2003). *Understanding violence* (2nd edn.). Mahwah, N.J.: Lawrence Erlbaum Publishers, Inc.

Mednick, S. A., Gabrielli, W. F. and Hutchings, B. (1984). Genetic influences in criminal convictions: evidence from an adoption cohort. *Science* **234**, 891–894.

Werner, E. E. and Smith, R. S. (1982). *Vulnerable but invincible: a study of resilient children*. New York: McGraw-Hill.

Viral Virulence and Stress

A J L Macario and E Conway de Macario
New York State Department of Health and The University at Albany (SUNY), Albany, New York, USA

Published by Elsevier Inc.

Scope
Viral Infection and ER Stress
Chain of Events and Its Consequences for the Virus and for the Infected Cell
The Endoplasmic Reticulum
ER Stress
ER Stress and Apoptosis
ER Stress and Viral Infection: Examples
Stress and Viral Reactivation

Glossary

Apoptosis	Programmed cell death that occurs in an orderly fashion, following a preexisting pathway and in response to specific inducers.
Caspases	A family of enzymes essential for apoptosis; these enzymes carry a cysteine that is key for cleaving proteins after an aspartic acid residue.

Endoplasmic reticulum (ER) An organelle of the eukaryotic cell formed by a single membrane, which is an extension of the inner membrane of the nuclear envelope. The ER membrane is folded onto itself around a space, the lumen, which is in direct communication with the intranuclear space.

ER stress Cellular stress at the level of the endoplasmic reticulum; may be caused by viruses.

Golgi apparatus An organelle of the eukaryotic cell present in plants and animals (although not in the cells of most fungi). Its chief functions are protein processing and sorting; also called the Golgi body or Golgi complex.

Mitochondrion A eukaryotic cell organelle present in most organisms. The mitochondrion's primary role is the conversion of fuel, in the form of organic materials, into energy, in the form of ATP. Mitochondria interact with the ER, particularly in situations leading to apoptosis.

Molecular chaperone A large group of special proteins that play an essential role in protein homeostasis. They are in charge of maintaining the entire complement of proteins in all cells and cellular compartments in their native functional conformations. The chaperones residing in the ER are particularly relevant to viral infection and its consequences for the cell through the induction of the unfolded protein response.

Necrosis Uncontrolled cell death; the consequence of acute, unanticipated cell or tissue injury.

Programmed cell death A controlled mechanism of cell elimination by suicide that multicellular organisms possess.

Signal transduction The cellular process of converting one kind of signal or stimulus into another. This process is usually very fast and involves a series of intracellular biochemical reactions in a specific sequence, signaling cascade, requiring specific enzymes and second messengers.

Unfolded protein response (UPR) A series of events that starts with the accumulation of unfolded, partially folded, misfolded, or aggregated proteins in the ER lumen, which triggers ER stress, inhibiting the functions of the organelle. This emergency situation then requires immediate action, which is mediated by the signal transduction cascade and results in the upregulation of some genes and in the downregulation of other genes.

Virulence The capacity of an infectious agent (e.g., a bacterium or a virus) to invade a host and cause lesions and disease.

Virus A biological particle that contains nucleic acid and other components, such as proteins and lipids, assembled in a shape that is typical for each viral species. A chief characteristic of viruses is that they require a host cell in which to replicate.

Many viruses cause cellular stress at the level of the ER (ER stress) and elicit the unfolded protein response (UPR). The main causative event is the accumulation of unfolded and/or misfolded proteins inside the ER; this accumulation immediately activates signal transduction pathways directed toward the nucleus, causing the upregulation of some genes and downregulation of others. This increases the proteins necessary to deal with stress in the ER while reducing many nonessential proteins, thereby lessening the task of processing them as it would under normal circumstances. If the ER burden is overwhelming, beyond mitigation, cell death occurs via apoptosis. Viruses control these processes to ensure that the cell will not die until they have replicated in sufficient numbers to achieve efficient propagation upon cell disintegration following apoptosis.

Scope

Viral infection has an impact on various biological functions at the level of the whole organism and at the levels of organs, tissues, cells, subcellular structures, molecules, and genes. Thus, the topic of viral virulence and stress can be treated from a variety of perspectives, centering on the immune system, the endocrine system, the gastrointestinal tract, the blood and blood-forming tissues, or many other systems and tissue and cell types. Moreover, the approach to the subject matter considered here can also be diverse, for example, clinical, histopathological, cytological, subcellular, and molecular. We focus on the cell of eukaryotic multicellular organisms (e.g., humans and rodents) and discuss the impact of viral infection on the ER, specifically virus-induced ER stress, the ensuing UPR, and the ER molecular chaperones.

Viral Infection and ER Stress

A number of viruses are known to cause ER stress and to induce the UPR (Table 1). These viruses are of interest because they are pathogenic to humans and to animals valuable to humans. Some of these viruses cause very serious diseases. For example, hepatitis C virus (HCV) causes hepatitis that can develop into liver cirrhosis and then into hepatocellular carcinoma. Others, such as the human cytomegalovirus (HCMV), cause mild illnesses in otherwise healthy individuals,

Table 1 Some viruses that cause ER stress and trigger the UPR[a]

Name	*Family*	*Subfamily*	*Genus*	*Group*	*Nucleic acid*	*Pathology*
HCV	Flaviviridae	–	*Hepacivirus*	IV	ssRNA(+)	Non-A non-B hepatitis, liver cirrosis, hepatocellular carcinoma
Bovine viral diarrhea virus	Flaviviridae	–	*Pestivirus*	IV	ssRNA(+)	Diarrhea in cattle
Japanese encephalitis virus	Flaviviridae	–	*Flavivirus*	IV	ssRNA(+)	Encephalitis
HCMV or HHV-5	Herpesviridae	β-Herpesvirinae	*Cytomegalovirus*	I	dsDNA	Infectious mononucleosis, retinitis, serious infections in people with defective or immature immune systems[b]
HHV-1/ HSV-1	Herpesviridae	α-Herpesvirinae	*Simplexvirus*	I	dsDNA	Oral and genital herpes
EBV or HHV-4	Herpesviridae	γ-Herpesvirinae	*Lymphocryptovirus*	I	dsDNA	Infectious mononucleosis, Burkitt's lymphoma
HRSV	Paramyxoviridae	Pneumovirinae	*Pneumovirus*	V	ssRNA(−)	Respiratory tract infections
TULV	Bunyaviridae	–	*Hantavirus*	V	ssRNA(−)	Hemorrhagic fever[c]
Murine retroviruses	Retroviridae	–	Several	IV	ssRNA-RT	Tumors and leukemias in mice

[a]ds, double-stranded; EBV, Epstein–Barr virus; ER, endoplasmic reticulum; HCMV, human cytomegalovirus; HCV, hepatitis C virus; HHV-1/HSV-1, herpes simplex virus 1; HHV-4, human herpes virus 4; HHV-5, human herpes virus 5; HRSV, human respiratory syncytial virus; ss, single-stranded; TULV, tula hantavirus; UPR, unfolded protein response.
[b]Fetuses, newborns, immunosuppressed individuals; and AIDS patients.
[c]Chiefly affected are the kidneys, lungs, and heart.

but they can cause very serious infections and even death in immunocompromised individuals (e.g., a patient who has received an organ transplant and is undergoing immunosuppressive therapy to prevent rejection of the transplanted organ). Still other viruses cause non-life-threatening but troublesome lesions; for example, herpes simplex virus 1 (HSV-1) produces painful blisters (herpes) of the oral cavity and lips.

Chain of Events and Its Consequences for the Virus and for the Infected Cell

It is important to realize that although a viral infection has a direct impact on the ER, it also affects other organelles, particularly the Golgi apparatus. Furthermore, it affects not only the originally infected cell but also many others in the vicinity and in tissues far away from the original infection site.

Another important consideration is that viral infection, viral virulence and its impact on the cell (the ER in particular), induction of the UPR and apoptosis, and cell destruction must all be regulated so the virus can multiply and propagate to other cells and tissues. If the damaging impact of the viral infection on the cell is too strong and too fast, the cell will die too soon for the virus to complete its cycle of infection and replication and the cell will disintegrate before a sufficient number of viral particles have been assembled for efficient propagation to other cells and organs via the blood.

Viruses need the host cell – its genes, proteins, and structures – for the replication and assembly of progeny virus. Thus, evolutionarily successful viruses have acquired genes and mechanisms that allow them to use the host cell's assets, even as they trigger ER stress and the UPR, en route to apoptosis. Some of the genes acquired by viruses for controlling the cellular response to infection encode anti-apoptotic proteins, which temporarily counteract the action of pro-apoptotic factors that are produced by the host cell on viral infection.

The Endoplasmic Reticulum

Subcompartments

Although the ER is a single organelle, present in one copy per cell, it includes specialized subcompartments, or domains, with specific functions.

- The rough ER (RER) is in contact with associated ribosomes, is close to the nucleus, and is involved in protein synthesis and in the translocation of newly synthesized polypeptides from the cytosol into the ER lumen.

- The transitional ER (tER) appears as cup-shaped cisternae and is involved in the exit of newly made polypeptides carried in coat protein complex II (COPII) vesicles toward the secretory route.
- The smooth ER (SER) participates in calcium storage, in lipid and steroid synthesis, and in the breakdown of drugs and chemical compounds, such as those present in medical remedies, which occurs predominantly in liver cells. The SER involved in calcium storage is prominent in muscle cells, including those of the heart and limbs; thus the SER plays a key role in muscular contraction, namely body movements and heart beats.

The Endoplasmic Reticulum is a Highly Dynamic Organelle

The ER varies quantitatively and qualitatively from cell to cell, depending on the functional status of the cell and on the tissue or organ in which the cell resides. Likewise, the distribution of the ER subcompartments inside a cell varies in relation to the cell-cycle stage and functional phase. For example, SER usually occurs in relatively small quantities, but its abundance is greater in cells that process and secrete lipids or steroids (e.g., in adrenal gland cells actively producing steroid hormones). RER is very abundant in cells that actively synthesize proteins, such as digestive enzyme-producing cells in the glandular tissue lining some portions of the gastrointestinal tract, and in plasma cells that actively produce immunoglobulins.

Specialized ER subcompartments also change within a cell, as illustrated by the switch from SER to RER, and *vice versa*; such a transition occurs in cells switching from low to high rates of protein synthesis, followed by return to the initial resting phase.

The Endoplasmic Reticulum and the Golgi Apparatus

Another important organelle of the eukaryotic cell is the Golgi apparatus, which interacts continuously with the ER. The Golgi is built of vesicles, tubules, and cisternae piled close to one another. All of these structures are formed by membranes composed of a huge range of proteins, perhaps well over a 1000 different varieties of proteins. The Golgi also possesses an enormous capacity for morphological transformation. This dynamism of the Golgi is affected by many factors, such as cellular environmental conditions, the cell's mitotic and physiological stages, the presence of drugs, stressors, disruption of microtubules, and protein movement from the ER.

The unusual plasticity of the Golgi membranes is probably due to the fact that proteins associate with the membrane only temporarily while they are traveling from the ER to other cellular locales. In fact, no specific class of protein has been identified that is uniquely and permanently part of the Golgi membranes. Golgi components are continuously entering and exiting this organelle.

The Golgi apparatus participates in protein glycosylation, cleavage of preproteins to produce the active form of the protein, and sorting of molecules so that they are properly tagged and thus directed to their physiological destinations, whether the cytosol, the cell membrane, or the extracellular space. The number of copies of the Golgi apparatus per cell varies from one to many, according to cell type, cell location (tissue type), and physiological and cell-cycle stage.

The Golgi can be subdivided into several regions, on the basis of function. The cis region lies near the ER and the nucleus and receives vesicles with their cargo of molecules coming from the SER. The middle region is the site of protein and lipid processing (e.g., glycosylation), and the trans zone or network region is sited near the cell membrane. In this region of the Golgi, preproteins are cleaved to generate functional protein molecules (as part of the maturation pathway from nascent polypeptide to fully processed, native protein) and the fully processed proteins are sorted according to their final destinations.

Interaction of the Endoplasmic Reticulum with Other Cellular Structures

The ER communicates directly with the nucleus, as mentioned earlier. Likewise, the ER (in the tER regions) interacts with or is modified by mitochondria, mRNAs, plasma membrane, peroxisomes, and phagosomes. These structures affect, from the outside, the shape and functionality of the ER, generating the various regions of this organelle, which suggest at first glance a variety of different ERs. Important within the context of this article is the interaction of the ER with mitochondria, a key stage in apoptosis.

Stress Proteins and Molecular Chaperones in the Endoplasmic Reticulum

A set of molecular chaperones resides and functions in the ER (Table 2). The ER chaperones participate in protein folding, refolding, and degradation, as do chaperones in other cellular compartments. However, the ER chaperones are involved as well in other processes that are unique to this organelle, such as calcium binding and protein translocation (exit from and entry into the ER lumen). It is of interest that some of the genes encoding ER chaperones are induced by glucose deprivation, which explains why these chaperones are classified and named as glucose-regulated proteins (e.g., Grp78, for glucose-regulated protein 78; see Table 2).

Table 2 Some stress proteins and chaperones found in the endoplasmic reticulum[a]

Mass (kDa)	*Examples*[b]
100 or higher	Hsp170 (Grp170)
81–99	Hsp90 (GRP94; Grp94; endoplasmin; gp96); PDIase: ERpdj5
65–80	BiP(Grp78); PDIase: ERp72, and ERp65; Kar2p; Lhs1p (Ssi1p, Cer1p)
55–64	RBP; PDIase: PDI, ERp57 (GRP58), PDIp, and PDIr; calnexin
35–54	Hsp40; Hsp47; Sec63p; ERdj3; Scj1p; JEM1p; PDIase: P5, ERp44, ERp46, TMX2long, TMX3, and TMX4; calreticulin
34 or lower	Cpn20; PDIase: ERp27, ERp28, ERp18, TMX, and TMX2short

[a]Stress proteins are those whose parent genes are induced by stress caused by any stressor. However, not all stress proteins are molecular chaperones; likewise, not all chaperones are stress proteins. Examples of chaperones residing in the ER are listed, whether or not they are stress proteins. There are also many other stress proteins in the ER (not listed here) that are not chaperones or that are not known at this time to function as chaperones.
[b]Synonyms are given in parentheses. Chaperones build multimolecular assemblies in order to exercise their functions. These assemblies are formed by identical or different chaperone molecules (subunits) and cofactors. In addition, these chaperone complexes interact with other chaperone complexes (networking) and with the protein-degradation machinery (e.g., the ubiquitin-proteasome system) to integrate the protein quality-control system of the cell.

The importance of the ER chaperones, both in normal physiological events and in dealing with stress generated by a variety of stressors (including viruses), emphasizes the critical, deleterious effects that defective chaperones have on the cell, and on the organism in general. A newly recognized type of pathology, the chaperonopathy, encompasses chaperones that are abnormal due to genetic (e.g., mutation) or acquired (e.g., aberrant posttranslational modification) defects and the impact of such abnormal chaperones on health and disease. The consequences of a chaperonopathy in the ER have not yet been evaluated, but they can be predicted to be significant for the cell and for any virus that causes ER stress. It is likely that chaperone molecules become deficient with age, due to the same biochemical modifications that affect other proteins in the elderly; thus, the probability that viral infections will cause serious ER stress and cell damage is likely to increase with senescence.

ER Stress

A variety of stressors, including certain viruses (Table 1), can cause ER stress and can trigger the ER-stress response. This variety of stressors notwithstanding, the common central factor underlying ER stress is the presence in the ER lumen of denatured proteins – partially or totally unfolded, or misfolded proteins. It is crucial for survival that, under ER-stress conditions, the cell readjusts a series of mechanisms that will reduce the number of ER tasks (essentially by diminishing the number of peptides that travel through and are modified in the ER) while increasing its functional power by expanding the ER's capacity to modify and tag proteins and to assist them in folding and refolding, thus steering them toward a productive pathway.

The ER plays a critical role in the monitoring of protein quality as the molecules traverse the organelle. It possesses the means to detect and to correct aberrant peptides and abnormal protein conformations. However, if correction is impossible, the ER has mechanisms to lead unrecoverable peptides toward degradation via the ER-associated degradation (ERAD) pathway. Finally, if all else fails, the ER can steer the cell toward apoptosis.

In addition to protein folding and modification (e.g., glycosylation), the synthesis of lipids and sterols occurs in the ER. Recent data suggest that lipid synthesis is altered as much as protein quality control during ER stress and the ensuing UPR.

A key event in the UPR is the increased expression of the genes encoding the ER-resident chaperones (Table 2). When the ER is overwhelmed by an excess of too many bad proteins, signals travel from the organelle to the nucleus, resulting in the upregulation of some genes and the downregulation of others. The downregulation reduces the nonessential biosynthetic tasks (e.g., tasks pertaining to the synthesis of secretory proteins) of the ER, which must redirect its resources to deal with stress. Still other mechanisms, such as the inhibition of mRNA translation, cooperate in reducing the synthesis of nonessential proteins.

ER Stress and Apoptosis

Apoptosis, a form of programmed cell death, is a multistage process that involves various types of molecules, many of them enzymes. The series of stages – the apoptotic cascade – can be induced in various manners that may schematically be organized into two main pathways: intrinsic and extrinsic. Intracellular injury (e.g., DNA damage and ER stress) starts the intrinsic or ER-stress apoptotic pathway, whereas extracellular agents activate the extrinsic (or receptor-mediated) pathway. In the latter pathway, the initial steps require the participation of cell-surface receptors to engage the enzyme caspases (cysteinyl aspartic acid proteases) and start the caspase cascade. This cascade includes the caspase-8-dependent cleavage of caspase-3 and the participation of caspase-8 in the cleavage of Bcl-2-associated

protein 31 (Bap31; Bcl stands for B-cell leukemia), and the recruitment of Bcl-2 proteins that interact with mitochondria and cause the leakage of cytochrome c from the organelle into the cytosol. This is followed by formation of a caspase-9 complex and apoptosis.

In contrast, the intrinsic pathway involves the mobilization of caspase-12 with the direct activation of caspase-9, initially bypassing the mitochondria. However, the intrinsic pathway also involves mitochondria (and cytochrome c leakage) through Bap31 cleavage with participation of caspase-12, resulting in the amplification of the caspase cascade.

Bcl-2 homologous antagonist/killer (Bak) and Bcl-2 associated X (Bax) proteins (i.e., two members of the pre-apoptotic class of proteins) become inserted into the ER membrane at the beginning of the ER-stress pathway, causing calcium efflux, which in turn causes an increase in cytosolic calcium. Calpain is activated, as is procaspase-12 in the ER. The caspase cascade proceeds up to the activation of procaspase-3. At this junction, there is no involvement yet of apoptotic protease-activating factor-1 (Apaf-1; a component of the apoptosome, whose main components are Apaf-1, cytochrome *c*/dATP, and procaspase-9), nor is there leakage of cytochrome c from the mitochondria into the cytosol. While the apoptotic cascade progresses, cytosolic calcium is taken up by the mitochondria as it escapes from the ER and mitochondrial membrane pores open, allowing cytochrome c to pass into the cytosol. At this point, the apoptosome is assembled and becomes active, as does procaspase-3.

The extrinsic route to apoptosis involves other factors: the c-Jun N-terminal kinase (JNK) inhibitory kinase (JIK), tumor necrosis factor receptor-associated factor 2 (TRAF2), and inositol-requiring enzyme 1α (Ire1α or IRE1α). A trimeric complex is assembled that includes IRE1α, TRAF2, and apoptosis-signal regulating kinase 1 (ASK1), resulting in the activation of ASK1 and JIK. This is followed by cell death.

The intrinsic and extrinsic pathways interact. TRAF2 promotes procaspase-12 aggregation and is released from proscapase-12 when ER stress occurs. This may happen because TRAF2 sequesters IRE1 and thus activates procaspase-12.

Some of the steps just described require additional experimental confirmation. Although these apoptotic pathways require caspases, suicide in some cells (e.g., certain neurons) does not involve these enzymes but resort to other mechanisms. In general, no matter which pathway is considered, caspase-dependent or -independent, the details of the mechanisms involved are still incompletely understood and may vary among cell types. However, the data available do indicate that the intrinsic and extrinsic apoptotic routes intersect and that the ER is at center stage in the programmed cell death triggered by ER stress.

ER Stress and Viral Infection: Examples

Hepatitis C Virus

HCV, a single-stranded RNA virus, causes non-A non-B hepatitis in humans. The potential consequences of the viral infection are the development of hepatic cirrhosis and eventually hepatocellular carcinoma. The ER is the site of HCV core processing and maturation; these processes involve the interaction of viral components with the ER membrane.

We have seen that the UPR is characterized by the activation of signals from the ER to the nucleus, resulting in the increased expression of ER-chaperone genes, calcium-binding proteins, and sarcoplasmic-ER calcium-transporting ATPase 2 (SERCA2), whose function is to pump calcium from the cytosol into the ER. In contrast, other proteins are decreased, via either the downregulation of their parent genes or the inhibition of the translation of their mRNAs. In this process, double-stranded-RNA-activated protein kinase-like ER kinase (PERK) plays a role, by mediating the phosphorylation of the eukaryotic initiation (of translation) factor 2α (eIF2α). When ER stress is intense and/or prolonged, apoptosis can occur in one of the following ways, depending on the animal species and other factors: (1) the activation, in mouse cells, of caspase-12 (not yet found in humans); (2) the activation of the C/EBP-homologous protein (CHOP; also known as the growth arrest- and DNA-damage-inducible gene 153, GADD153); or (3) the induction of JNK, involving a complex formed by TRAF2 and ASK1.

Calcium signaling is also dependent on ER function. The ER controls the opening of the inositol 1,4,5-triphosphate receptor (IP3R) channels and calcium uptake into the lumen by the SERCAs. The mitochondrion also participates in calcium trafficking; this organelle is very efficient in calcium uptake and release. Some mitochondria are situated very close to the ER and thus are conveniently positioned to sense changes in calcium concentration in the cytosol as luminal calcium leaks out of the stressed ER through activated calcium (IP3R) channels.

The HCV core induces the expression of the chaperones BiP(Grp78), Hsp90(Grp94), and calreticulin (Table 2) and SERCA2 at levels well above those characteristic of the unstressed ER. The increase in the levels of all these proteins is enhanced and paralleled by the loss of calcium from the ER lumen. Eventually, the HCV core triggers the apoptotic cascade due to the hyperexpression of CHOP/GADD153

(which acts as a pro-apoptotic factor), Bax translocation to mitochondria, mitochondrial membrane depolarization, cytochrome c efflux, and caspase-3 activator and poly-(ADP-ribose) polymerase (PARP) cleavage.

Human Cytomegalovirus

HCMV belongs to the Herpesviridae family, along with the varicella-zoster virus, HSV-1, and HSV-2, which cause chickenpox and shingles, and Epstein–Barr virus (Table 1). Epstein–Barr virus and HCMV are the most common etiological agents of infectious mononucleosis. HCMV remains dormant in infected cells without generating symptoms, but it can be activated and can generate serious cell injury in individuals with immature immune systems (e.g., fetuses, newborns, and patients with congenital deficiencies), and in patients with impaired immunity, such as those being treated with drugs that kill lymphocytes and other blood cells (e.g., cancer patients undergoing chemotherapy and individuals who have received an organ transplant and are being given immunosuppresants to prevent rejection of the foreign tissue).

HCMV induces ER stress, which is sensed by means of activating transcription factor 6 (ATF6), PERK, and IRE1. Normally, in the absence of stress, BiP(Grp78) (see Table 2) is bound to each of the sensors on the luminal side of the ER membrane. If unfolded, misfolded, or aggregated proteins accumulate in the ER, BiP(Grp78) releases the sensors and interacts with the proteins to assist them to fold correctly so that they can regain their native functional conformation. When PERK and IRE1 are no longer bound to BiP(Grp78), their luminal domains form homodimers and become phosphorylated and activated. Simultaneously, ATF6, also freed from the chaperone, exposes a Golgi localization signal on its structure and is thus translocated to the Golgi, in which it is cleaved and activated.

PERK activation initiates a pathway for the attenuation of translation that involves eIF2α, eIF2β, activating transcription factor 4 (ATF4), and other gene proteins. At the same time, ATF6 mediates transcriptional upregulation of ER chaperone genes. ATF6 is an ∼50-kDa protein produced by the digestion of a 90-kDa preprotein in the Golgi. As a result of the ATF6-initiated pathway, the levels of ER chaperones increase in response to the increased demand due to accumulation of client polypeptides in the ER lumen.

Also part of the virus-initiated UPR is the induction of the IRE1 pathway, increasing the expression of genes whose protein products are involved in promoting protein degradation. In this manner, defective peptides that cannot be refolded by the chaperones are eliminated via digestion, and the digestion-derived amino acids are reused on the ribosome to synthesize essential proteins (e.g., chaperones) as required by the increased demand due to ER stress.

In summary, HCMV intervenes in the regulation of the three routes of signal transduction characteristic of the UPR, the routes mediated by PERK, IRE1, and ATF6. The HCMV intervention ensures that the virus has time to replicate while the cell is preparing for apoptosis, which leads to cell disruption and the release of viral particles. In this regard, it has been found that the genes encoding ER chaperones are also induced by the virus independently of the ATF6-initiated pathway, possibly because the chaperones are needed for the generation of proteins required by the virus. In parallel, the protein degradation ER degradation-enhancing 1,2-mannosidase-like (EDEM) factor was found to be inhibited, so as to transitorily curtail proteolysis and thus ensure the presence of the molecules necessary for cell survival during viral replication.

Stress and Viral Reactivation

Viral infection causes cell stress, as just explained. The reverse, the stress activation of latent viruses, also occurs. Some viruses become latent after the acute phase of infection; their presence in the cell causes no clinically detectable effects. The viral genome remains in the infected cell but is transcriptionally repressed except for a minor region; hence the virus is inactive or latent. A number of cell stressors, such as fever, ultraviolet light, and social tensions, with their accompanying neuroendocrine imbalance, induce the reactivation of these latent viruses. Viruses belonging to the Herpesviridae family (Table 1), (e.g., HSV-1, varicella zoster virus, and Epstein–Barr virus) are among those that pass through active and latent stages depending on the microenvironmental cellular conditions. Upon the impact of stressors, the mechanisms that keep the viral genome transcriptionally repressed are disrupted and the virus begins to replicate and accumulates in the host cell, usually generating pathological and clinical consequences. For instance, when HSV-1, a neurotrophic virus (it invades neurons), is reactivated in ganglionic cells, viral particles are produced, which emerge from the host cell and invade other cells that thus suffer viral-induced stress. The molecular mechanisms that control latency and reactivation are poorly understood. Likewise, the molecular mechanisms that mediate stress-induced viral reactivation have not been fully elucidated and the role of molecular chaperones in these reactivating mechanisms remains to be determined.

See Also the Following Articles

Chaperone Proteins and Chaperonopathies; Heat Resistance; Heat Shock Genes, Human; Heat Shock Response, Overview; Herpesviruses; Chaperonopathies; Proteases in the Eukaryotic Cell Cytosol; Proteases in Prokaryotes and Eukaryotic Cell Organelles.

Further Reading

Benali-Furet, N. L., Chami, M., Houel, L., et al. (2005). Hepatitis C virus triggers apoptosis in liver cells by inducing ER stress and ER calcium depletion. *Oncogene* **24**, 4921–4933.

Ciechanover, A. (2005). Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nature Reviews Molecular Cell Biology* **6**, 79–86.

Efstathiou, S. and Preston, C. M. (2005). Towards an understanding of the molecular basis of herpes simplex virus latency. *Virus Research* **111**, 108–119.

Ellgaard, L. and Ruddock, L. W. (2005). The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Reports* **6**, 28–32.

Federovitch, C. M., Ron, D. and Hampton, R. Y. (2005). The dynamic ER: experimental approaches and current questions. *Current Opinion in Cell Biology* **17**, 409–414.

Isler, J. A., Skalet, A. H. and Alwine, J. C. (2005). Human cytomegalovirus infection activates and regulates the unfolded protein response. *Journal of Virology* **79**, 6890–6899.

Kleizen, B. and Braakman, I. (2004). Protein folding and quality control in the endoplasmic reticulum. *Current Opinion in Cell Biology* **16**, 343–349.

Kroemer, G. and Martin, S. J. (2005). Caspase-independent cell death. *Nature Medicine* **11**, 725–730.

Lee, A. S. (2001). The glucose-regulated proteins: stress induction and clinical applications. *Trends in Biochemical Sciences* **26**, 504–510.

Levine, T. and Rabouille, C. (2005). Endoplasmic reticulum: one continuous network compartmentalized by extrinsic cues. *Current Opinion in Cell Biology* **17**, 362–368.

Li, X-D., Lankinen, H., Putkuri, N., et al. (2005). Tula hantavirus triggers pro-apoptotic signals of ER stress in Vero E6 cells. *Virology* **333**, 180–189.

Macario, A. J. L. and Conway de Macario, E. (2004). The pathology of anti-stress mechanisms: a new frontier. *Stress* **7**, 243–249.

Macario, A. J. L. and Conway de Macario, E. (2005). Sick chaperones and cellular stress: pathological and clinical implications. *New England Journal of Medicine* **353**, 1489–1501.

Macario, A. J. L., Grippo, T. M. and Conway de Macario, E. (2005). Genetic disorders involving molecular-chaperone genes: a perspective. *Genetics in Medicine* **7**, 3–12.

Miller, C. S., Danaher, R. J. and Jacob, J. (2006). ICP0 is not required for efficient stress-induced reactivation of herpes simplex virus type 1 from cultured quiescently infected neuronal cells. *Journal of Virology* **80**, 3360–3368.

Munro, S. (2005). The Golgi apparatus: defining the identity of Golgi membranes. *Current Opinion in Cell Biology* **17**, 395–401.

Panet, A., Braun, E., Honigman, A., et al. (2005). Involvement of cellular death signals in the reactivation of herpes simplex virus type 1 and lambda bacteriophage from latent state. *Journal of Theoretical Biology* **236**, 88–94.

Puthenveedu, M. A. and Linstedt, A. D. (2005). Subcompartmentalizing the Golgi apparatus. *Current Opinion in Cell Biology* **17**, 369–375.

Schroeder, M. and Kaufman, R. J. (2005). The mammalian unfolded protein response. *Annual Review of Biochemistry* **74**, 739–789.

Su, H-L., Liao, C-L. and Lin, Y-L. (2002). Japanese encephalitis virus infection initiates endoplasmic reticulum stress and an unfolded protein response. *Journal of Virology* **76**, 4162–4171.

Westhoff, B., Chapple, J. P., van der Spuy, J., et al. (2005). HSJ1 is a neuronal shuttling factor for the sorting of chaperone clients to the proteasome. *Current Biology* **15**, 1058–1064.

Zhang, X., Uthaisang, W., Hu, L-F., et al. (2005). Epstein-Barr virus-encoded latent membrane protein 1 promotes stress-induced apoptosis upstream of caspase-2-dependent mitochondrial perturbation. *International Journal of Cancer* **111**, 397–405.

Viruses and Stress

G Z Feuerstein and R R Ruffolo
Wyeth Research, Collegeville, PA, USA
B-N David
Israel Institute for Biological Research, Ness-Ziona, Israel

© 2007 Elsevier Inc. All rights reserved.

Introduction
Stress: Viral Infection and Penetration to the CNS
Cold Stress Neuroinvasiveness and Virulence of Attenuated WNV
Summary

Glossary

Attenuated West Nile Virus (WNV)	A virus that has lost its neuroinvasiveness but not its neurovirulence.
Blood–brain barrier (BBB)	Property of endothelial cell brain microvessels that provides a barrier for transport of molecules from the systemic circulation into the brain. The BBB is considered to have an important function that provides the central nervous system with a constant environment needed for normal brain function. Stress conditions disrupt the BBB permeability, which is often increased in many stress situations.
Glucocorticosteroids (GCS)	Organic lipid molecules produced primarily in the adrenal gland. GCS serve as local and systemic hormones and act on specific receptors that activate multiple genes and signaling pathways governing multitude biological actions. Most prominent in GCS action are glucose, protein, and lipid metabolism, regulation of immune and inflammatory cells and mediators, bone and muscle remodeling, as well as central nervous system effects.
Neuroinvasiveness	Ability to penetrate the central nervous system.
Neurovirulence	Ability to establish a lethal infection within the central nervous system.
Plaque-forming unit (PFU)	A standard measure of viral load.
Viral encephalitis	Viral infection of the brain. In order to cause encephalitis when invading the central nervous system, the virus must possess neurovirulence.
Virulence	Ability of a virus to cause clinical disease and potential for mortality. Wild-type West Nile virus is both neurovirulent and neuroinvasive.
West Nile Virus (WNV)	The etiologic agent of West Nile fever, an endemic viral disease in many subtropic regions and an emerging mosquito-borne disease in the United States. WNV is a neuroinvasive flavivirus that may cause systemic and central nervous system infection, paralysis, and death.

Introduction

Stress is defined as any external or internal stimuli that disrupt the natural homeostasis of the organism. Stress situations commonly result in suppression of the immune system. The reaction of the organism to stress can be adaptational and helpful in restoring optimal homeostasis to maintain the organism functions or can lead to a variety of aberrant emotional and physical conditions with long-lasting consequences.

The effect of stress on the humoral immune response has been reported in numerous studies. The responses to stress vary greatly depending on the stress patterns (type and duration) and the type of stressor. The effect of stress on the immune system in human and animal studies is invariably suppressive, on both the cellular and humoral levels.

Physiological responses to stressors are largely mediated by the sympathetic nervous system and the hypothalamic-pituitary-adrenal (HPA) axis. Glucocorticoids (GCS) are major mediators in the reaction to stress and in stress-induced immunosuppression.

Stress stimuli affect the immune system and cause immunosuppression through involution of lymphoid organs, possibly mediated by increased activity of the pituitary-adrenocortical axis. There is also a clear connection between stress, the central nervous system (CNS), and the immune response. The role of the CNS in immune responses to stress conditions associated with infectious viral conditions is also well established. Glucocorticoids released in response to stress are the major mediators of the effect of stress on viral replication, change in virulence, and viral penetration to the CNS.

Stress: Viral Infection and Penetration to the CNS

A variety of stress paradigms have been shown to exacerbate the effects of several infectious agents, including herpes simplex virus, coxsackie B1 virus,

vesicular stomatitis virus, cytomegalovirus (CMV), influenza virus, encephalitic viruses, and mengovirus. Viruses are known to flourish in the stressed host, with increased morbidity and mortality. It appears that stress acts by enhancing virus replication in lymphocytes as shown in the spleen of stressed mice. Since activated lymphocytes can pass through the blood–brain barrier (BBB), the larger fraction of infected lymphocytes in stressed mice could also contribute to higher virus titer in the brain. This possibility, however, awaits further clarification by more direct evidence on the mode of immune cell penetration into the CNS.

An attenuated strain of West Nile virus (WNV) in mice had little capacity to induce encephalitis and death, as the level of the virus in the brain was very low. Specifically, brain virus levels in control mice inoculated with WN-25, a noninvasive variant of the West Nile virus, or SVN, a neurovirulent, noninvasive Sindbis virus, showed titers of less than 2 log 10 plaque-forming units (PFU), whereas the virus titers in the brain of cold-stressed mice or mice exposed to GCS treatment were 8.5, 7.8, and 7.5 (log 10 PFU) for WN-25 and 7.3, 7.1, and 7.0 (log 10 PFU) for SVN, respectively. In stressed mice, these viruses (WN-25 and SVN), both lacking neuroinvasive properties in control nonstressed mice, readily invaded the brain and proliferated to high titer levels (>106 PFU/brain) in all moribund mice inoculated with WN-25 or with SVN, on top of a stress condition. It was suggested that the neurovirulent properties of non-neuroinvasive viruses can serve as a marker for BBB permeability and aid in studying pathological processes of the brain.

Cold Stress Neuroinvasiveness and Virulence of Attenuated WNV

In mice inoculated with the wild-type strain of WNV, it was shown that short-term exposure of the mice to cold water ($1 \pm 0.5^{\circ}C$) for 8–10 days increased the mortality from 47% in control to 92–100% in the cold-stressed mice. Cold or isolation stress increased blood, spleen, and brain virus levels by two to three orders (100- to 1000-fold) as compared to control nonstressed mice. WNV, which was isolated from the brain of moribund stressed mice, was more virulent when inoculated intraperitoneally to normal nonstressed mice. Concomitantly, lymphoid organs such as spleen diminished in mass in the stressed mice. These observations were also demonstrated with WN-25 and SVN.

Since GCS levels are elevated during stress, the effects of stress or corticosterone (a rodent GCS) injection were investigated on neurovirulence and encephalitis by the attenuated arboviruses. These two viral strains were exposed to the same cold stress paradigm described previously. The stress paradigms induced mortality of 65, 80, and 50% (WN-25) and 75, 78, and 58% (SVN) in cold, isolation, and corticosterone-treated mice, respectively. No mortality was seen in nonstressed virus-inoculated mice. The level of WN-25 virus in the brain of stressed mice was 7.5–8.5 log 10 PFU. No virus was detected in nonstressed inoculated mice. Isolation of a neuroinvasive strain from brains of stress-exposed mice suggested that the induction of infection was attributed to immunosuppression with subsequent increased viral replication and selection of neuroinvasive strain. The WN-25 virus extracted from the brain of moribund stressed mice was extremely virulent; inoculation of as little as 5 PFU to normal nonstressed mice caused severe encephalitis and death. WN-25 showed a change in its neuroinvasive properties, raising the possibility of stress-enhanced proliferation and a selective process leading to reversal of the neurovirulence. In contrast, brain SVN virus levels isolated from moribund mice (stress exposed) were no different from levels of viruses in the nonstressed mice, suggesting a different role of stress in the mechanism of stress-induced virulence.

The susceptibility of stressed mice to the viral infection could be explained by reduced number and suppression of macrophage and T-lymphocyte activity as well as by decreased natural killer cell cytotoxicity. Therefore, stress can induce elevated viremia for longer periods, facilitating the penetration of viruses into the CNS.

Summary

There is strong evidence for an association between stress and the immune system and between the immune system and the CNS. Susceptibility to infections is presumed to be primarily mediated by immune function. Since WNV is widely distributed and is capable of endemic spread, the effect of stress on WNV infection has been studied. It seems that suppressed immune response during stress, with defective function of macrophages and lymphocytes, leads to enhancement of viral replication in the blood and peripheral organs, induction of a selective process, and selection of a neuroinvasive strain.

See Also the Following Articles

Glucocorticoids, Role in Stress; Immune Suppression; Infection; Viral Virulence and Stress.

Further Reading

Ben-Nathan, D. (1994). Stress and infectious disease. *Israel Journal of Veterinary Medicine* **49**, 105–112.

Ben-Nathan, D. and Feuerstein, G. (1990). The influence of cold or isolation stress on resistance of mice to West Nile Virus. *Experientia* **46**, 285–290.

Ben-Nathan, D., Lustig, S. and Feuerstein, G. (1989). The influence of cold or isolation stress on neuroinvasiveness and virulence of an attenuated variant of West Nile Virus. *Archives of Virology* **109**, 1–10.

Ben-Nathan, D., Lustig, S. and Kobiler, D. (1996). Cold stress-induced neuroinvasiveness of attenuated arboviruses is not solely mediated by corticosterone. *Archives of Virology* **141**, 1221–1229.

Ben-Nathan, D., Kobiler, D., Loria, R. M. and Lustig, S. (1998). Stress-induced CNS penetration by non-invasive attenuated encephalitis viruses. In: Levy, A., Grauer, E., Ben-Nathan, D. & de Kloet, E. R. (eds.) *New frontiers in stress research: modulation of brain function*, pp. 277–283. Amsterdam: Harwood Academic Publishers.

Cohen, S. and Williamson, G. M. (1991). Stress and infectious disease in humans. *Psychological Bulletin* **109**, 5–24.

Dhabhar, F. S., Miller, A. H., McEwen, B. S. and Spencer, R. L. (1995). Effect of stress on immune cell distribution. *Journal of Immunology* **154**, 5511–5527.

Isler, J. A., Skalet, A. H. and Alwine, J. C. (2005). Human cytomegalovirus infection activates and regulates the unfolded protein response. *Journal of Virology* **79**, 6890–6899.

Lustig, S., Olshevsky, U., Ben-Nathan, D., Lachmi, B. E., Malkinson, M., Kobiler, D. and Halevy, M. (2000). A live attenuated West Nile strain as a potential veterinary vaccine. *Viral Immunology* **13**, 401–410.

Mehta, S. K., Cohrs, R. J., Forghani, B., Zerbe, G., Gilden, D. H. and Pierson, D. L. (2004). Stress-induced subclinical reactivation of varicella zoster virus in astronauts. *Journal of Medical Virology* **72**, 174–179.

Sheridan, J. F., Dobbs, C., Brown, D. and Zwilling, B. (1994). Psychoneuroimmunology: stress effects on pathogenesis and immunity during infection. *Clinical Microbiology Reviews* **7**, 200–212.

Welsh, C. J., Bustamante, L., Nayak, M., Welsh, T. H., Dean, D. D. and Meagher, M. W. (2004). The effects of restraint stress on the neuropathogenesis of Theiler's virus infection II: NK cell function and cytokine levels in acute disease. *Brain Behavior and Immunity* **18**, 166–174.

Vomeronasal Organ *See: Pheromones.*

W

Waist–Hip Ratio

R Rosmond
Björndammsterrassen 41, Partille, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by P Björntorp and R Rosmond, volume 3, pp 669–671,
© 2000, Elsevier Inc.

Waist–Hip Ratio
Waist–Hip Ratio as a Risk Factor
Visceral Adipose Tissue
Visceral Obesity
Mechanisms of Waist–Hip Ratio as a Risk Factor
Waist–Hip Ratio as a Sign of Stress

Glossary

Hypothalamic-pituitary-adrenal (HPA) axis	The hormonostatic endocrine axis that regulates cortisol production and secretion.
Lipoprotein lipase	The rate-limiting enzyme that cleaves one fatty acid from a triacylglycerol.

Various techniques are available for the assessment of body composition. These techniques comprise use of anthropometric measurements, radiographic techniques, and more complex techniques such as total body potassium and tetrapolar bioelectrical impedance analysis.

The double thickness of the epidermis and subcutaneous fat is determined by a skin-fold measurement. The skin-fold thickness can be measured wherever a fold can be formed, but specific sites such as the triceps and subscapular locations have been designated because of their high correlations with total body fat.

Body circumferences, like skin folds, can be measured at various body locations. The sites for circumference measurements that have been reported to be important in describing the amount of abdominal, visceral obesity are the waist and hip circumferences and their ratio, waist–hip ratio (WHR). In addition, the WHR has been reported to be a powerful predictor of obesity-related diseases such as cardiovascular disease, diabetes mellitus, hypertension, and hyperlipidemia. Although the abdominal fat distribution pattern is more common in men, both men and women show increased risk of serious disease with greater abdominal fat. Men may be considered at increased risk if the WHR is greater than 0.95, and women if the WHR is greater than 0.80.

Cortisol regulates adipose-tissue differentiation, function, and distribution, and in excess, causes visceral obesity. The presence of a continuously changing and sometimes threatening external environment may, when the challenges exceed a threshold, activate the hypothalamic-pituitary-adrenal (HPA) axis. The resulting increase in circulating cortisol directs body fat to visceral depots, mainly because of the high density of glucocorticoid receptors in visceral adipose tissue. Thus, central obesity may be regarded as a sign of frequent repeated or chronic perceived stress of various origins, including socioeconomic and psychosocial disabilities, alcohol, tobacco, depression, and anxiety.

Waist–Hip Ratio

Body circumferences and their ratios are used to indicate the distribution of body fat. The most important variations, in terms of health associations, are between the amounts of fat in internal, mainly intra-abdominal sites, as distinct from subcutaneous sites. The ratio of waist–hip circumferences was the first anthropometric method developed from epidemiological research as an indicator of fat distribution in relation to metabolic diseases. Most of the value in indicating body fat is derived from waist circumference, the hip circumference probably reflecting several other body tissues such as bones and muscles.

Waist–Hip Ratio as a Risk Factor

There is a well-documented sex dimorphism in regional adipose tissue distribution. Indeed, despite the fact that women are usually more obese as a group than men, male subjects more frequently have significantly higher mean waist circumference and higher mean WHR in agreement with the greater propensity of men to accumulate excess fat within the abdominal cavity. Since the WHR has been shown to be associated, albeit moderately, with the amount of abdominal visceral adipose tissue measured by computed tomography or magnetic resonance imaging, this index has been widely used to investigate the relations between regional adipose tissue distribution and metabolic profile. Thus it was effective in predicting aberrations in glucose and insulin levels and also showed a strong correlation between plasma lipids and blood pressure. WHR predicted subsequent diabetes in men and coronary heart disease in both men and women, and was more predictive of these endpoints than either the body mass index ($kg\ m^{-2}$) or a more complex procedure using the sum of multiple skin-fold thicknesses.

Visceral Adipose Tissue

Adipose tissue is specialized connective tissue that functions as the major storage site for fat in the form of triglycerides. Visceral adipose tissue has a high cellular density as compared to subcutaneous fat. In addition, the blood flow and innervation are more abundant than in subcutaneous adipose tissue. This makes the visceral adipose tissue highly metabolically active, with a turnover of fat that is considerably higher than in other fat depots. This is exaggerated by the high density of specific receptors for steroid hormones, which permit a powerful effect of triggering hormones for adipocyte metabolism, such as catecholamines. Insulin plays a predominant role in the lipogenic process. The net effect of insulin is to enhance storage and block mobilization and oxidation of fatty acids. Insulin exerts its effect by stimulating the formation of lipoprotein lipase (LPL), so that circulating triglycerides are hydrolyzed and free fatty acids can enter adipocytes. Insulin is also required for the transport of glucose, which is needed for re-esterification of triglycerides once inside adipocytes. Finally, the conversion of glucose to fatty acids is accomplished by activation of several enzymes by insulin. Cortisol in the presence of insulin stimulates, via the high density of glucocorticoid receptor (GR) in visceral adipocytes, a lively expression of LPL by transcription of the LPL gene and posttranslational stabilization of the enzyme. Cortisol and insulin exert powerful effects to both promote lipid accumulation in adipocytes and diminish lipid mobilization in general. This is most pronounced in visceral adipocytes due to the higher density of GRs.

In concert with growth hormone, sex steroid hormones have opposite effects. Testosterone in the presence of growth hormone affects the stimulation of lipid mobilization, exaggerated by the effect of testosterone to upregulate its own receptor density. This is accomplished by transcriptional effects at several levels of the lipolytic cascade of events, including the density of β-adrenoceptors, cyclase, and hormone-sensitive lipase. Simultaneously, the LPL activity is inhibited. Estrogen has also been found to inhibit LPL.

Visceral Obesity

As mentioned above, several of the hormones of the endocrine system are involved in controlling the rate and direction of fat metabolism. While cortisol and insulin promotes accumulation of visceral fat, growth hormone and sex steroids, often in concert, result in the opposite. Cortisol also exerts chronic effects on lipid metabolism. One of the most striking in humans is the centralization of body fat seen in Cushing's syndrome, exhibited as severe truncal obesity. After successful therapy, the somatic features of Cushing's syndrome disappear.

Mechanisms of Waist–Hip Ratio as a Risk Factor

Visceral fat seems to be metabolically unique compared with subcutaneous abdominal fat. Dr Björntorp suggested that visceral fat results in hepatic insulin resistance via a portal effect of free fatty acids released by increased omental fat. The increased flux of free fatty acids to the liver leads to increased hepatic glucose production and decreased hepatic insulin clearance, which in turn leads to powerful risk factors for diabetes, cardiovascular disease, and stroke. The visceral accumulation of depot fat, as indicated by the simple WHR index, is thus an easily visible and measurable consequence of a profound, multiple endocrine perturbation.

Waist–Hip Ratio as a Sign of Stress

There can be little doubt that chronically elevated cortisol levels can redistribute body fat, as clearly seen in Cushing's syndrome. Excessive and protracted cortisol secretion is the result of a frequently evoked HPA axis, either by environmental or intrinsic factors. Environmental factors include psychosocial and socioeconomic handicaps, alcohol, and tobacco, as well as traits of psychiatric symptoms of depression and anxiety. Such stressors are all known to activate the HPA axis. In primates other than humans,

exposure to stress is followed by a diminished feedback regulation of the cortisol secretion as well as visceral obesity and associated metabolic complications. The WHR may therefore be regarded as a simple anthropometric measurement of frequently repeated or chronic exposure to various forms of stress.

Further Reading

Björntorp, P. (1993). Visceral obesity: a "civilization syndrome" *Obesity Research* **1**, 206–222.

Lean, M. J. and Han, T. S. (2002). Waist worries. *American Journal of Clinical Nutrition* **76**, 699–700.

Rosmond, R., Dallman, M. F. and Björntorp, P. (1998). Stress-related cortisol secretion in men: relationships with abdominal obesity and endocrine, metabolic and hemodynamic abnormalities. *Journal of Clinical Endocrinology Metabolism* **83**, 1853–1859.

Rosmond, R., Wallerius, S., Wagner, P., et al. (2003). A 5-year follow-up study of disease incidence in men with abnormal hormone pattern. *Journal of Internal Medicine* **254**, 386–390.

War Stress in the Former Yugoslavia

M Flögel, S Šupraha Goreta and G Lauc
University of Zagreb, Zagreb, Croatia

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by M Flögel and G Lauc, volume 3, pp 678–683, © 2000, Elsevier Inc.

Historical and Political Background of the Conflict
Migrations, Ethnic Cleansing, and Social Impact
Effects on Health
Effects on Mortality
Effects on Social Behavior and Violence
Conclusion

Glossary

Posttraumatic stress disorder (PTSD) — PTSD is a psychological disorder affecting individuals who have experienced or witnessed profoundly traumatic events, such as torture, murder, rape, or wartime combat, characterized by recurrent flashbacks of the traumatic event, nightmares, irritability, anxiety, fatigue, forgetfulness, and social withdrawal.

War is the multifaceted threat to human existence. It exposes people to the health-damaging psychological, physical, chemical, biological, and mechanical stressors. The war has a deep and devastating social and economic impact on the population involved, thus producing additional and lasting repression. It is too soon to make long-term estimates of the damages caused by war in the area of former Yugoslavia, but the data concerning increased social and health impairments due to the war in Croatia are indicative for the whole area and show the level and the type of vulnerability when humans are exposed to the stress of war. War integrates the deprivation of basic existential needs and of all human rights and values; it eliminates emotional comfort; it causes irreversible material and kin losses, physical exhaustion, and psychological breakdowns; and it makes futile all everyday routines. The first and the only benign response to stress is the escape, but there is no possibility of running away from stress experiences in war – one stressor can only be substituted by another. This is why the full display of adaptation mechanisms takes place in subjects experiencing war stress: changes in behavior, hormonal impairments, neurological disorders, metabolic changes, genetic shifts, and all kinds of diseases. Stress-induced consequences, depending on the severity and the duration of the exposure to the stress experience, can be identified as PTSD, interrupted pregnancies, susceptibility to malignant diseases, cardiovascular insults, suppressed immune response, hormonal disorders, and all types of emotional instabilities or psychological impairments including suicide, homicide, and other types of violence.

Historical and Political Background of the Conflict

The federal republics of prewar Yugoslavia had national and cultural identities with distinct historical backgrounds. At times in the past they were sovereign states, and they have been trying to reestablish their independence ever since they lost it. Neither the first (monarchist) nor the second (communist) Yugoslavia was an expression of the people's will. When the

Berlin wall collapsed, the Soviet block was dismantled, the Soviet Union disintegrated, and the ruling communist party was defeated. The sweeping changes in eastern Europe provided an opportunity for the national republics constituting Yugoslavia to secede freely, as the Czech and Slovak republics have done. In response, the Yugoslav army under the command of the Serbian generals tried to annul the free and democratic decision by the electorate in four of the six Yugoslav republics to separate from Yugoslavia. The Serbs, inspired by the idea of Greater Serbia, were the only nation to support the continuation of Yugoslavia. The Serbian idea of turning Yugoslavia into a greater Serbia triggered the outburst of the latent but strong mistrust and intolerance between the Serbs, on one side, and all other nations, on the other. The Communist Party of Serbia, subsequently re-formed as the Socialist party, had control over the powerful Yugoslav army and turned into an aggressive Greater Serbian force that started to conquer and ethnically cleanse as much as possible of the territory of the republics whose majority of the population had voted for independence. The local people, if not ethnic Serbs, had to leave the occupied territories. Deprived of all goods and rights, they experienced the misery of homeless and jobless refugees. Those who tried to defend their homes were either massacred and killed or put into concentration camps, tortured, and raped. The process started in Croatia in 1991 and ended by North Atlantic Treaty Organization (NATO) intervention in Kosovo in 1999. Instead of democratization and prosperity, first the Croats, then the Bosnians, and at the end the Albanians experienced war, violence, fear, exile, and loss of kin. Macedonia and Montenegro experienced the menace of social and economic turbulence introduced by the hundreds of thousands of refugees that poured into their countries. Only the Slovenians succeeded in withstanding the austerities of Serbian aggression. At least 2.5 million victims (including almost the total population of Kosovo) were expelled from their homes and were, for varying periods, in exile or temporary housing.

Migrations, Ethnic Cleansing, and Social Impact

The first victims of the Serbian aggression were the Croats, who were removed from 30% of Croatian territory. That territory was not only occupied by the Yugoslav army and the local Serbs and ethnically cleansed of Croats; in addition, Serbs from other parts of Croatia or Bosnia were persuaded to resettle in abandoned Croatian homes in the territory cleansed for Greater Serbia. Such behavior escalated national animosities as the aggression spread further and further. In Bosnia and Herzegovina, 70% of the territory was ethnically cleansed of Croats and Bosnians, and only the Dayton Agreement and NATO intervention stopped the war in Bosnia and forced the Serbs to restrict their control to the Serbian Republic (Republika Srpska), encompassing 49% of Bosnian territory.

In 1991, Bosnia had 4 364 600 inhabitants; 250 000 men were killed during the war and 35 000 women were raped. At the peak of the war, almost 2 million people were displaced. Even a decade after the end of the war in Bosnia, several hundred thousand people still had not returned to their homes.

The total population of Croatia (in 1991) was 4 760 000. It is estimated that 260 000 people left the country during the war. The war in Croatia killed 14 000 people, among whom 43.4% were civilians and 213 were children; 34 000 veterans, 9890 adult civilians, and 1142 children were injured; 7551 people were detainees of concentration camps and 23 000 children lost one or both parents. In addition, 389 433 civilians were displaced within the country (it is estimated that this number is 6% larger because of unregistered people). More than 100 000 houses were destroyed. Hospitals, schools, and many plants and farms were demolished, drastically raising the unemployed and retired percentages of the population and putting exhausting pressure on the remaining working population. There were 870 000 retired people, 260 000 jobless, and only 970 000 employed (compared to 1 650 000 employed before the war). The collapse of order and the rise of poverty, insecurity, and hopelessness were widely reported throughout the population.

Effects on Health

One of the major problems in the interpreting of the effects of war on the health of the population is the great extent of migration associated with the war. Croatia is the only country in the region where health parameters can be followed before, during, and after the war. Although migrations from Croatia and within Croatia were large, the number of displaced people never exceeded 12% of the total population and all the victims of migrations within Croatia are included in the official statistics.

The statistical parameters depict a very gruesome picture of the effects of war on the health of the population in Croatia. The rates of several diseases that are known to be associated with stress, such as diabetes, tumors, cardiovascular diseases, and gastric ulcers, increased sharply during and immediately after the war (Figures 1–2). Infectious diseases also increased

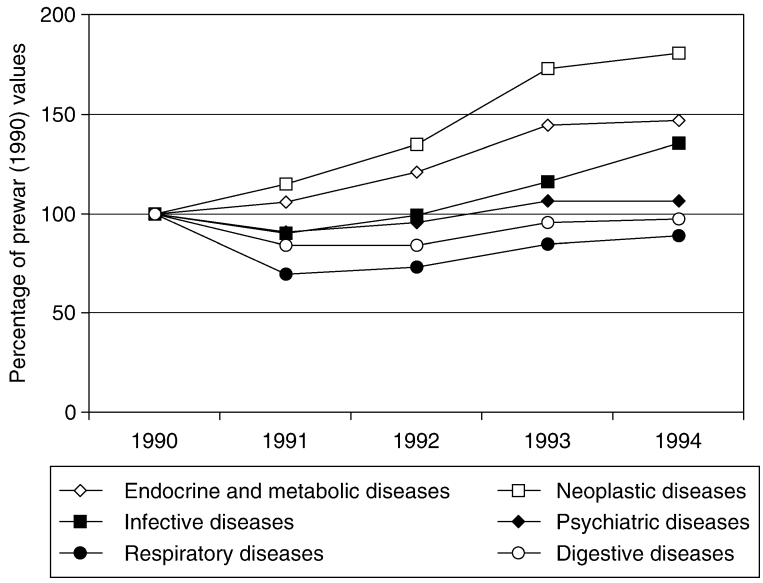


Figure 1 Number of cases of various disease groups registered in general hospitals in Croatia, 1990–1994. Data from Croatian Health Service Yearbook 1997, Zagreb, 1998.

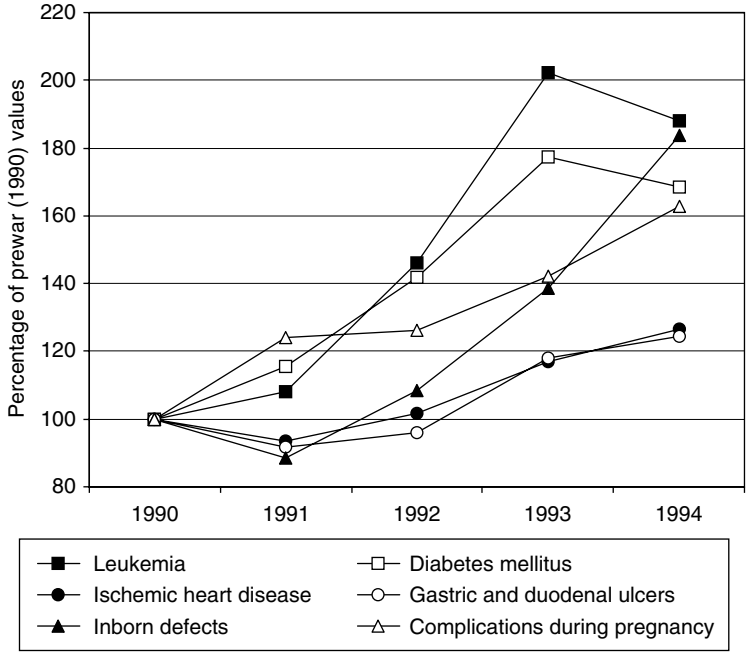


Figure 2 Number of cases of specific diseases registered in general hospitals in Croatia, 1990–1994. Only diseases with the highest increase of incidence are presented. Data from Croatian Health Service Yearbook 1997, Zagreb, 1998.

significantly, probably due to the stress-induced depression of the immune system. Most other diseases displayed a slight decrease in 1991–1992 (mirroring the temporary population decrease at the peak of the war in Croatia) and subsequently returned to the prewar values. The exception is rheumatoid arthritis, which stayed significantly decreased for several years and then increased rapidly to over 150% of prewar values. During this time, infant mortality continually declined, indicating that there was no significant decrease in the general quality of the health-care system. Despite the preserved level of health care, the postwar health of the population significantly deteriorated due to stress and other effects of war compared to 1990. In 1994, Croatia had 60 000 more cases of infective diseases, 40 000 more people with diabetes mellitus, 32 000 more people with various neoplastic diseases, 30 000 more patients with heart diseases, 40 000 more people

with kidney diseases, and 24 000 additional patients with duodenal or gastric ulcers. Because the total population of Croatia was approximately 4.5 million, the addition of 200 000 ill people represented a substantial burden for an already overloaded health system.

Stress due to the war had a particularly strong impact on pregnancy and prenatal development. In addition to a decreased birthrate, as is usual in the time of hardship, there was a continuous increase in extrauterine pregnancies and other complications during pregnancy and delivery. In addition and completely unexpected, there was a sharp increase in various inborn defects (Figure 2).

Effects on Mortality

In addition to the decreased birthrate, the war diminished the population by increasing the mortality (Figure 3). Not taking into account the 14 000 people killed during the war, the most prominent immediate effect of war on the Croatian territory was the more than doubling of the rate of mortality from ischemic heart diseases, resulting in the unexpected deaths of more than 6000 people during the two war years 1991 and 1992. The prolonged effects of war, manifested in a ruined economy and prewar social structures, still kills people with an annual rate of approximately 4000. The increase in the ischemic heart disease alone has taken more human lives than were killed during the whole war in Croatia. Unfortunately, due to the massive migrations in Bosnia and Kosovo, it is not possible to determine the exact numbers for them, but if the percentage of people killed was similar to that in Croatia, several hundred thousands people probably died because of stress.

It is worth noting that, in addition to the stress imposed by the war, the transition from a centralized socialistic economy to a free-market economy introduced stressful existential problems to a part of the population, which probably also contributed to the increased mortality from ischemic heart disease. One recent study in Lithuania reported that the breakup of the Soviet Union resulted in a quadruple increase in the mortality rate from ischemic heart diseases for some age groups. However, in other countries, such as Hungary and Poland, this pattern is not so prominent. The transition in Croatia is expected to be less stressful than in other socialist countries because even during socialism Croatia was open to the West and had accepted a few elements of the market economy. Thus, the war and its consequences remain the major factors influencing the increased mortality from ischemic heart disease in Croatia.

Effects on Social Behavior and Violence

Quantitative analysis of the psychological effects of war on the Croatian population is limited to the subjects who were/are seeking medical help. However, only a small number of individuals suffering from psychological disorders are under medical surveillance. There are more then 10 500 diagnosed cases of PTSD; many more are unregistered. Studies estimated that patients suffering from PTSD have up to a sevenfold increased incidence of suicides and fourfold increased risk of death from all external sources. Violence in general has increased, but it is very hard to know the exact figures because the health-monitoring statistics do not include violent behavior in their reports and official policy makers in Croatia refuse to acknowledge the incidence of violence and suicides to be disturbing. However, some trends are clearly visible. During and immediately after the war (1991–1992), there was a twofold increase in the level of violence, both as reported in police statistics and in hospital admissions. Although, according to

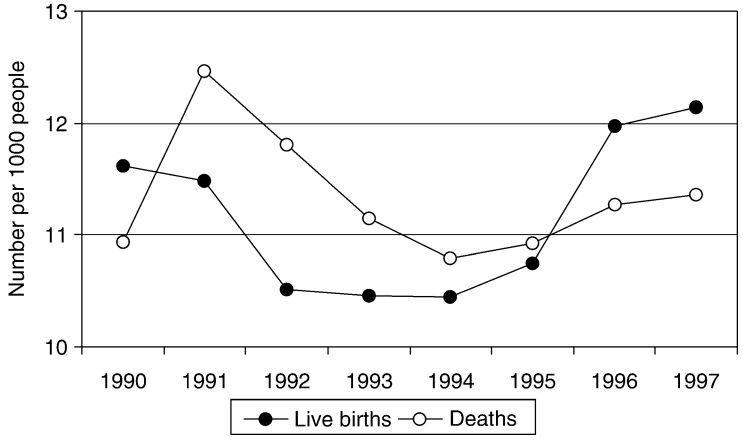


Figure 3 Number of live births and deaths per 1000 people, 1990–1997. Data from Statistical Yearbook 1998, Republic of Croatia Central Bureau of Statistics, 1998.

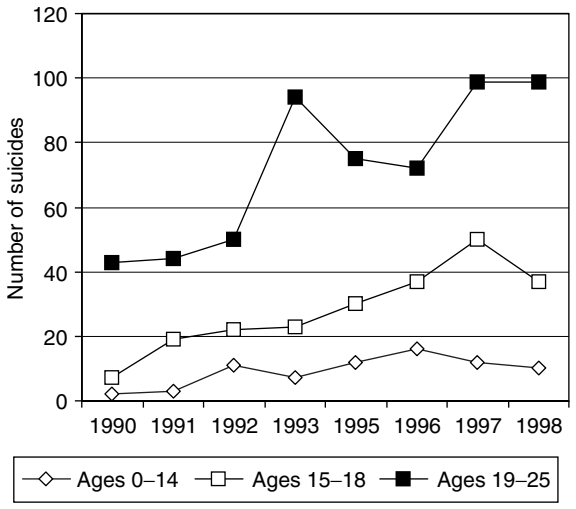


Figure 4 Number of suicides and attempted suicides in different age groups in Croatia, 1990–1998. Data from official statistics of the Ministry of Interior of the Republic of Croatia.

the same sources, the reported homicides and violent crime returned to prewar values in subsequent years, there is a disquieting increase in suicides and other self-inflicted injuries in young people. The number of young people committing suicide is continuously rising, and compared to 1990, in 1997 this rate is doubled or quadrupled, depending on the age group (Figure 4). In the last four years, primary and secondary schools report a rapidly growing number of hyperactive and aggressive children who need special care in the classroom. The incidence of depressive symptoms and insomnia is reported to be great in the elderly and in veterans. Drug abuse is disturbingly frequent compared to the prewar statistics, but the number of treated addicts has remained constant. It is correct to attribute these changes in behavior to the stress imposed by the war, and it is interesting how the various target groups (children, the elderly, veterans, etc.) within the population have responded differently to the stress.

Conclusion

The majority of the population in each of the states originating from the former Yugoslavia has experienced war stress. Various health and social disorders evolved from that experience. Although the data presenting the consequences of the war stress are incomplete and although they are, because of the availability, reduced to Croatia only, they provide unique evidence of human vulnerability when living through war.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Disasters and Mass Violence, Public, Effects of; Gulf War Syndrome, Psychological and Chemical Stressors; Nuclear Warfare, Threat of; Persian Gulf War, Stress Effects of; Posttraumatic Stress Disorder, Delayed; Vietnam Veterans, Postwar Experiences and Health Outcomes; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Clinical.

Further Reading

Babic-Banaszak, A., Kovacic, L., Kovacevic, L., et al. (2002). Impact of war on health related quality of life in Croatia: population study. *Croatian Medical Journal* **43**, 396–402.

Black, P. H. and Garbutt, L. D. (2002). Stress, inflammation and cardiovascular disease. *Journal of Psychosomatic Research* **52**, 1–23.

Breslau, N., Kessler, R. C., Chilcoat, H. D., et al. (1998). Trauma and posttraumatic stress disorder in the community: the 1996 Detroit Area Survey of Trauma. *Archives of General Psychiatry* **55**, 626–632.

Bullman, T. A. and Kang, H. K. (1994). Posttraumatic stress disorder and the risk of traumatic deaths among Vietnam veterans. *Journal of Nervous and Mental Disease* **182**, 604–610.

Dekaris, D., Sabioncello, A., Mazuran, R., et al. (1993). Multiple changes of immunologic parameters in prisoners of war: assessments after release from a camp in Manjaca, Bosnia. *Journal of the American Medical Association* **270**, 595–599.

Kocijan-Hercigonja, D., Sabioncello, A., Rijavec, M., et al. (1996). Psychological condition hormone levels in war trauma. *Journal of Psychiatric Research* **30**, 391–399.

Lauc, G., Zvonar, K., Vukšić-Mihaljević, Ž., et al. (2004). Post-awakening changes in salivary cortisol in veterans with and without PTSD. *Stress Health* **20**, 99–102.

Lukenda, J., Kolaric, B., Kolcic, I., et al. (2005). Cardiovascular diseases in Croatia and other transitional countries: comparative study of publications, clinical interventions, and burden of disease. *Croatian Medical Journal* **46**, 865–874.

Markotic, A. (2002). Balkan syndrome. *Lancet* **359**, 166.

Marsland, A. L., Bachen, E. A., Cohen, S., et al. (2002). Stress, immune reactivity and susceptibility to infectious disease. *Physiology & Behavior* **77**, 711–716.

Reiche, E. M., Nunes, S. O. and Morimoto, H. K. (2004). Stress, depression, the immune system, and cancer. *Lancet Oncology* **5**, 617–625.

Sabioncello, A., Kocijan-Hercigonja, D., Rabatic, S., et al. (2000). Immune, endocrine, and psychological responses in civilians displaced by war. *Psychosomatic Medicine* **62**, 502–508.

Sapolsky, R. M. (1996). Why stress is bad for your brain. *Science* **273**, 749–750.

War, Suicide and Sacrifice

V Hazboun
Guidance and Training Center for the Child and Family, Bethlehem, Palestine

© 2007 Elsevier Inc. All rights reserved.

Background
Theory
Clinical Research
Clinical Experience
Research
Stress Management in War

Glossary

Intifada	Palestinian uprising against the occupying forces.
Nakbeh	The disaster (Arabic).
Sacrifice	The slaughter of an animal or a person or the surrender of a possession as offering to a deity, made holy by religious association.
Suicide	The act of a person intentionally killing him- or herself.
War	A quarrel between nations conducted by force, a state of open hostility.

Background

It is the hope of the author to present, from a clinical perspective, the experience of war, suicide, and sacrifice. For centuries, people have launched wars with other nations and peoples to pursue ideals, goals, or possessions. Others have fought wars in self-defense and in anticipated concern about their safety. Others have been invaded and occupied and have resisted the occupation through warlike tactics. People have committed suicide, for a purpose or for no purpose, because of an ideal or an illness. People have sacrificed so that others can benefit; this altruism has been a characteristic that elevates humans from the selfish and decadent position that they can become fixated on.

In the Middle East, the Palestinian-Israeli conflict goes back several decades, and it is impossible to address all the issues here. However, a summarized orientation to the problem is necessary to understand the trauma and stress sustained by this population. The following years are significant:

- 1948 – The Nakbeh (the disaster, in Arabic) is the word used by the Palestinian population to refer to the partitioning of Palestine and the massive Palestinian refugee exodus and displacement. The Israelis refer to this same period as their independence day.
- 1967 – The 7-Day War and the occupation of the rest of Palestine by the Israeli armed forces.
- 1989 – The first Intifada, the Palestinian people's uprising against the occupying forces.
- 1994 – After the Oslo talks, the declaration of the Palestinian Authority.
- 2000 – The second Intifada, a rekindling of the first Intifada after major disappointments and tragedies that continued in the area despite the Palestinian Authority's presence.
- 2004 – The building of the separation wall, still in progress in 2006.

Theory

The psychoanalytic model is a scientific perspective that has been used by the author in her 25 years of clinical experience as a psychiatrist dealing with traumatized patients. The author's clinical experience has allowed her the opportunity to compare cases from economically impoverished areas such as the ghetto of Los Angeles and the refugee camps in Palestine, along with more economically viable areas such as Hollywood and Jerusalem. War here is considered to be a state of open hostility possibly causing trauma to individuals and the masses alike. The individual or national reaction to such hostility remains a very personal experience, even from a national perspective. Some people are traumatized from such hostility; others are rendered stronger and more resilient, determined, and creative. Both the aggressors and those experiencing aggression should reflect on this and perhaps learn to refrain from aggressive behavior; aggressors give those experiencing aggression a chance to become noble and most creative, resilient, and determined, and if those experiencing aggression become aggressive in return they lose their values and become the same as the original aggressors.

Trauma is an experience that can be caused from within or from the outside. That means that people can experience trauma internally without anyone contributing to this phenomena. This situation of internal conflict is well known to psychiatrists, who are constantly challenged with such circumstances. External traumatizing events and their impact on people's psychological well-being have been very clearly and concisely defined and presented in the diagnostic criteria for posttraumatic stress disorder

(PTSD) in the *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.; DSM-IV) and subsequent articles on the subject.

The Middle East conflict seems to tie together many important aspects of our civilization: politics and the search for democracy; religion or the lack of it, and its abuse through fanaticism; justice and its continued absence, leading to intolerable injustice; violence and aggression; gender difference; equality and the lack of it; pride and the feeling of superiority and being chosen, leading to prejudice and the dehumanizing of the other; occupation that is viewed as liberation for others; terrorizing others and the feeling of being terrorized at the same time; and war and the yearning for peace. Here I present a psychoanalytic look at the Middle East conflict, hoping that it sheds light on the 50-year conflict and encourages the different parties to engage in deep reflection, seeking the answers in an arena that has not been explored before. I believe that human nature and people's unconscious processes and defenses are interfering with the capacity of the people involved to seek and realize peace, whereas a scientific understanding of these processes might alter the escalating cycles of violence and convince people to adopt a different approach, one that will lead to a lasting peace. This then could serve as a model for other areas of escalating conflict and confrontation.

Let us first examine the basic ideas behind the understanding that led to the conclusions presented here. The work of Sigmund Freud was at the basis of the understanding that was later developed by Melanie Klein that the power of the unconscious and its impact on human existence needs to be addressed. To become conscious of unconscious processes liberates people, who can then either choose to be led by their neurotic thoughts or choose to venture into a life of free choice, risk taking, and responsible action led by thought and reflection rather than by impulse and magical thinking. Freud taught us about the defense mechanisms in the human mind and attitudes, namely the early defenses of splitting and projecting. As evidenced in many human relationships, and certainly in politics, the good and the bad reflect this very notion of splitting, seeing the self as good and the other as the enemy on whom all kinds of evil projections can be accommodated. Melanie Klein further developed these thoughts with the idea of projective identification. She suggested that when magical thinking is used to develop a person's wishful thoughts (in action or in thought) it renders the person responsible for the possible damage that occurs to other people, regardless of the real cause of this damage. The sense of responsibility or guilt may lead to self-punitive behavior or feelings, but it is the feeling of entrapment that can cause more damage, tarnishing the person's life; that is, the person may feel that, if he or she magically caused irreversible damage in other people's lives, then this magical fate is the only possible fate.

Clinical Research

The work presented here is based on many years of clinical observations.

- 1984 – The author observed and analyzed the drawings of 50 children in three refugee camps in Palestine. The drawings were compared to those of children in the same age group in the Los Angeles County Hospital's children's inpatient unit. The drawings of the children living in the Palestinian refugee camps were more often about healthy community activities scenes such as picking olives in the fields, weddings, and home scenes; the inpatient children's drawings in Los Angeles were more morbid, showing bars, blood, and death scenes and issues.
- 1991–1994 – As mental health supervisor for the United Nations Relief and Works Agency (UNRWA), the author gathered clinical evaluations. The reports indicated that, of 1500 new cases seen yearly, 450 children were diagnosed as having pervasive developmental disorder, 350 as having depression, and approximately 250 as having PTSD.
- 1992 – During the first Intifada, the author visited the same three refugee camps as in the 1984 study and examined the drawings of 50 children in each camp. These drawings had an extremely different tone; they were overwhelming in their gory depictions of pain, blood, barbed wire, prison and death, violence, and injustice.
- 1996 – Mental health workers, social workers, and psychologists randomly administered 360 questionnaires in the West Bank. These questionnaires were translations from English of DSM-IV criteria. The diagnostic criteria for three disorders was analyzed and used: PTSD, depression, and pervasive developmental disorder. The results indicated an alarming 35%, 40%, and 45% incidence, respectively. These results will be reassessed soon to evaluate the impact of the second Intifada and the continued unrelenting trauma that has haunted the area.

Clinical Experience

It has been the experience of the author that, although war is traumatizing, it does not necessarily have a paralyzing effect on the population. Aside from the obvious trauma of the loss of people's lives and property, heritage and development, health and hope, it is the observation of the author after 20 years of living

among and serving this population that the reality of becoming refugees, being displaced, losing loved ones, and tolerating a great deal of frustration and humiliation has not turned the masses into desperate people who let their impulses and instincts control them; instead, people who recognize the wisdom of maintaining educational pursuits and developing future generations who will honor their traditions seem to be the norm. The fact that some people resort to the idealization of martyrdom, death, and violent suicide/homicide remains to be understood and investigated. Next I present the results of my clinical experience and research done to try to analyze and understand this phenomena.

As part of the initial assessment of the children at the Guidance and Training Center for the Child and Family (GTC) in Bethlehem, the author asks the children to express three wishes. In the mid-1980s, the children expressed many very creative wishes expressing a great many ambitions and objectives in life. In the mid-1990s, 10 years later, the majority of the children did not express any wishes; they remained silent. After some psychotherapeutic intervention, if they expressed any wishes, the wishes all could be summarized as a wish to become a *shaheed*, a martyr. This inability to express wishes followed by a wish to become a martyr indicates a paralysis of the mind and creativity and a total embracing of a death wish. These children can relinquish these thoughts after a solid psychotherapeutic intervention and a facilitation of their reintegration into the school system.

Clinical Vignettes

Patient 1 A young woman, a would-be suicide bomber, spoke on local public TV about her previous intentions to blow herself up for what she thought was a nationalistic pursuit and about how, feeling uncomfortable with her plans, she sought psychotherapy. Then, she spoke about her change of heart after her psychotherapeutic experience at our center. However, she reported that she had been unable to convince her brother to come in for treatment and that he ended up as a suicide bomber, killing himself and others.

Patient 2 Another patient, a 7-year-old boy, was brought in for treatment because he was sleepwalking at night; this behavior endangered his life, especially because the city was under curfew during most of the period when the child exhibited this behavior. It turned out that this child had been arrested and taken for interrogation for throwing stones at the occupying forces. During the arrest, his grandmother tried to stop the soldiers from taking the child; she was pushed and she fell, hurting her head and landing in a pool of blood. The boy thought she was dead as he was taken for his interrogation. But on his release a few hours later, he saw that his grandmother was well and not dead as he had feared. Following this episode, the child started to exhibit a phobia about leaving the home and refused to go to school; he eventually presented with sleepwalking, which led to his being brought in for treatment.

During his therapy, when the boy was asked about dreams, he reported having a repetitive dream of someone kidnapping him and putting him in dangerous situations. When the author said that she might know who the kidnapper was, the child was surprised and interested in the answer. The author told him that she thought he was both the kidnapper and the kidnapped because he was the author of the dream and therefore was trying to tell us something about himself – that he kidnapped himself at night while asleep and put himself in dangerous situations. The dream also was residual trauma from the arrest and interrogation that he sustained, but it was interpreted at a deeper level because the child was receptive to this level of interpretation, and he eventually benefited from the interpretation. The sleepwalking stopped, and the child was able to go to school and did not continue to repeat the trauma that he had sustained by having this disturbing dream and exhibiting dangerous behavior that could have cost him his life. The child felt guilty about his grandmother's injury and felt he had caused it by his behavior, which led to the army's intrusion into their home. He felt identified with this projection and trauma and wanted to feel in control through this repetitive dream of the kidnapper. He felt guilty and entrapped by the fate that he thought he caused his grandmother to have.

Patient 3 Another patient, a breast-feeding mother, told the author, "This infant I can feed this way, but the others at home, how will I feed them? If someone were to give me a bomb, I would go to the checkpoint and blow myself up. This way, my children at least will be proud of their mother, the martyr." The author took this into consideration as she continued to treat this patient and offered her an income-generating opportunity at the clinic doing some embroidery. This woman is now taking care of her family and running an income-generating embroidery project of her own, helping other women sustain their livelihoods and those of their families.

Research

A brief study was carried out at the center to assess the impact of the separating wall being built by the Israeli government on the population living in the West Bank,

asking people's views of this new cement barrier separating them from the families that had been living next door to them prior to this construction. The study also assessed people's views vis-à-vis the Holocaust.

The questionnaire was given to 30 Palestinian professionals; the results are shown in Table 1. Twenty years ago, the results for the last four questions were as listed in Table 2. When the present-day results are compared to the assessments made 20 years ago, it becomes apparent that this construction has decreased the compassion people feel toward the other side and increased their feelings of being oppressed and persecuted.

After the construction of the separating wall and the population's increasing feelings of frustration and desperation, and the claim by the other side that the construction of the separating wall would provide security for the Israelis, the author investigated the opinion of the community about suicide bombers. The idea for this questionnaire came from the observation, first with children and then with adults, that those who came to the center for psychiatric and psychological treatment expressed wishes or plans to become suicide bombers. After these patients received psychological interventions at the center, which either facilitated their being mainstreamed back into the school system or into the workforce, these thoughts of suicidal/homicidal plans disappeared and were replaced with hope and gratitude toward the people who helped them find hope for the present and future; they relinquished their death wish and embraced a life wish.

The questionnaire was given to 91 mental health workers; the results are shown in Table 3. They indicate that a high percentage of the population honor and glorify this sacrifice and do not view it as reflection of emotional difficulties, seeing suicide bombing as the patriotic sacrifice of a hero, done in the name of God, and not as an expression of suicidal thoughts, depression, or psychological injury. The results show that suicide bombers are seen as desiring revenge, being aware of what they are doing, and not feeling guilty about the pain of other people. These ideas are not the same as those expressed by the population that comes for consultations at the guidance and training center, as presented in the earlier material.

Further research is planned that may help us further understand the state of mind of the populations in this area. This includes a study comparing trauma

Table 1 Assessment of the views of Palestinian professionals vis-à-vis the Holocaust and the separation wall being built by the Israeli government[a]

Question	*Yes*	*No*
1. Wall will destroy life in Palestine	28	2
2. Wall will help Palestinians get a Palestinian State	7	23
3. Wall will eliminate Palestinians by Israel	25	5
4. Israelis are doing to Palestinians what Hitler did to them	23	7
5. Consider Jews living through a massacre in Germany	28	2
6. Feel sad about Holocaust	22	8
7. Feel happy about the Holocaust	1	29

[a]*N* = 30, 20 males and 10 females.

Table 2 Assessment of the views of Palestinian professionals (questions 4–7) 20 years ago[a]

Question	*Yes*	*No*
4. Israelis are doing to Palestinians what Hitler did to them	20	10
5. Consider Jews living through a massacre in Germany	29	1
6. Feel sad about Holocaust	29	1
7. Feel happy about the Holocaust	1	29

[a]*N* = 30.

Table 3 Assessment of views about marytrdom (%)[a]

Do you think that the people who conduct martyrdom operations:	*Yes*	*No*	*No answer*
1. Desire to commit suicide?	17.6	76.9	5.5
2. Do a patriotic sacrifice?	84.6	12.1	3.3
3. Do a sacrifice in the name of God?	83.5	13.2	3.3
4. Do it due to their difficulties in life?	60.4	35.2	4.4
5. Suffer from depression?	49.5	44.0	6.6
6. Are psychologically hurt?	64.8	28.6	6.6
7. Desire revenge?	84.6	9.9	5.5
8. Are aware of what they are going to do?	45.1	49.5	5.5
9. Feel guilty about the pains of other people?	34.1	60.4	5.5
10. Are heroes?	46.2	49.5	4.4

[a]*N* = 91 mental health workers.

in Palestinian and Israeli populations and the capacity to forgive and a study comparing the mental health disorders observed in Palestinians living in West Bank to disorders in Palestinians living in Lebanon.

Stress Management in War

Although it is well known that suicide attempts, psychotic episodes, and admissions to psychiatric hospitals increase after a war is over and there is peace, it is the author's observation that the mental health demands in Palestine are increasing and becoming overwhelming. It is with this in mind that the interventions discussed here are being provided. The prediction that things will become even worse when there is peace in this area makes the author concerned about meeting these increasing mental health needs, and the author hopes that the prediction is wrong. The mental health services are not sufficient at the present time, and should things continue to deteriorate, the outcome is of concern to many health workers.

The use of psychoanalytic interventions and understanding, behavioral and relaxation techniques, special education therapy, individual (talk or play) and group therapy, medications, and other modalities to reduce stress and trauma in the population has become common. A thorough initial evaluation and treatment plan are necessary for each patient who consults a treatment center, and the provision of mental health services is urgent and necessary at this time.

That a majority of the population does not pursue psychotherapeutic assistance and yet seems to manage and maintain its existence and sanity implies that human nature is capable of self-healing and reasoning, of developing the necessary wisdom through people's own thinking and their culture and religion to pursue healthy avenues in their decision making, taming their aggressive impulses and discovering on their own that in this way creativity is allowed to develop and flourish. The few individuals pursuing violent resistance seem to hurt the cause and the people themselves more than help them. This leaves the author concerned about the repercussions of such attitudes; two wrongs never made a right. This is explained by Klein's projective identification defense, as previously elaborated on. The fact that war has contributed to the confusion of suicide with sacrifice should not cause mental health workers to confuse the issues; they should always try to sort things out in a clear and ethical way, trying to provide patients with inner harmony, helping them think more clearly and wisely through the psychotherapeutic intervention.

It is important to also mention that in times of war and political chaos and injustice it is the moral obligation of the mental health workers to maintain a standard of ethical service to the community and to try to attend to the needs of the masses by training others who can provide triage and quality interventions and by using the media (magazines, newspapers, radio, and television) to present mental health issues and prevention-measures training. All these different interventions have proved to be worthwhile efforts during the author's years of experience in the Middle East. Along with attending to the needs of the masses from a mental health perspective, it is also important to keep in mind that help along the lines of mainstreaming kids back into the school system is essential and that providing income-generating projects for the population may play a role in preventing disasters.

Working in a vacuum is not a creative way to intervene, and the involvement of professionals who are eclectic, versatile, creative, and passionate about their work makes the difference between touching people and helping them versus observing, reporting, and being afraid to get involved in the struggle to find and spread inner harmony.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Gulf War Syndrome, Psychological and Chemical Stressors; Persian Gulf War, Stress Effects of; Posttraumatic Stress Disorder in Children; Posttraumatic Stress Disorder, Delayed; Posttraumatic Stress Disorder, Neurobiology of; Vietnam Veterans, Postwar Experiences and Health Outcomes; War Stress in the Former Yugoslavia; War-Related Posttraumatic Stress Disorder, Treatment of; Posttraumatic Stress Disorder – Clinical; Posttraumatic Stress Disorder – Neurobiological basis for.

Further Reading

Adelson, E. (1975). *Sexuality and psychoanalysis*. New York: Brunner, Mazel Inc.

Freud, S. (1993). *The interpretation of dreams*. U.S.A: Basic Books.

Goldstein, J., Freud, A., Solnit, A. J., et al. (1986). *In the best interest of the child: the least detrimental alternative*. New York: The Free Press.

Hazboun, V. (2002). Autobiographical chapter. In: Armbruster, D. (ed.) *Tears in the Holy Land*. Wilsonville, OR: Bookpartners.

Hazboun, V. (2003). A psychotherapeutic view of violence. In: *Special issue on two traumatized societies. Palestine-Israel Journal of Politics, Economics and Culture* **10**(4) [online].

Hazboun, V. (2003). A psychotherapeutic view of violence and its alternatives. *RAHAT Medical Journal* **1**(5), 17–23.

Hazboun, V. (2004). La testimonianza di una psichiatra suelle conseguenze del lungo conflitto Palestinese. *Il Regno* **49**(945).
Hazboun, V. (2004). Psychotherapeutic interventions with Palestinian children and their families. *Dynamic Psychiatry* **204–205**, 69–77.
Hazboun, V. (2005). Clinical experience in the field of war and PTSD. *Trauma e Guerra, Rivista Sperimentale di Freniatria* **129**(1), 115–124.
Hazboun, V. (2005). L'intifada e la condizione dell'infanzia in Palestina: Nel Suo Nome, conflitti, riconoscimento, convivenza delle religioni. *Il Regno* 141–145.
Hazboun, V. and Ziskoven, M. (2002). Ein schmerzhafter frieden ist besser als die qualen des krieges. *Forum der Kinder-und Jugendpsychiatrie und Psychotherapie* **12**(2).
Hazboun, V. and Ziskoven, M. (2002). Ein schmerzhafter frieden ist besser als die qualen des krieges [A painful peace is better than agonies of war]. *Psychosoziale Arbeit in Bethlehem Gesundheitsmagazin*. Dr. Med. Mabuse, 136, Maerz.
Klein, M. (1984). *Love, guilt and reparation*. New York: The Free Press.
Lewis, M. (1982). *Clinical aspects of childhood development*. Philadelphia: Lea & Febiger.
Mason, A. (1983). The suffocating super ego. In: Grotstein, J. (ed.) *Dare I disturb the universe*. Caesura Press.
Segal, H. (1978). *Introduction to the work of Melanie Klein*. Hogarth Press.
Victorian Foundation for Survivors of Torture (1996). *A guide to working with young people who are refugees*. Victoria, Australia: Victorian Foundation for Survivors of Torture.
Wener, C. (1982). *Psychopathology from infancy through adolescence*. New York: Random House.

War-Related Posttraumatic Stress Disorder, Treatment of

L B Slone and M J Friedman
National Center for PTSD, U.S. Department of Veterans Affairs and Dartmouth Medical School, White River Junction, VT, USA

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by R Rosenheck and A Fontana, volume 3, pp 672–677, © 2000, Elsevier Inc.

Introduction
Stressors
Assessment and Diagnosis
Treatments
Summary

Glossary

Comorbid	Occurring at the same time as another medical condition.
Imaginal exposure	Exposure involving imagination, images, or imagery of the traumatic event.
In vivo exposure	Exposure to a real-life situation.
Pharmacological	Having to do with drugs.
Psychometric	Quantitative assessment of psychological trends and measures.
Psychotherapeutic	Of, relating to, or used in treatment of mental or emotional disorder or maladjustment involving verbal communication.
Reservists	Military personnel who serve part-time unless activated. Includes both National Guard and Reserve components.

Introduction

Some service members handle their transition from the war zone to the home front in a matter of weeks. Others require more time, and perhaps assistance. A significant minority fall short and develop posttraumatic stress disorder (PTSD), depression, alcoholism, some other psychiatric problem, or functional impairment. Although the present focus is PTSD, these other potential postdeployment problems should not be overlooked. The treatment of PTSD in veterans is similar in many ways to the treatment of PTSD in the civilian population, with some special considerations. Although there are universal aspects to any combat experience, every war is different, thereby posing unique challenges to the practitioner. For example, 50% of American troops that served in Iraq and have left active duty are members of the National Guard or Military Reserve rather than active duty combatants. This has implications for the identification and treatment of PTSD in this cohort.

Stressors

Traditional severe combat stress involves situations in which a service member experiences comrades being killed or injured; having personally killed enemy combatants and, possibly, innocent bystanders; uncontrollable and unpredictable life-threatening attacks; and more. In addition to direct combat stress, a variety of other war zone stressors are experienced, such as feeling helpless to alter the course of potentially lethal events, observing or handling the remains of civilians and soldiers, and observing devastated communities produced by combat. Military sexual trauma (MST) is now known to be highly prevalent, affecting both women and men in combat. An unprecedented number of military personnel are surviving serious injuries due to dramatic advances in medical technology, and being wounded in battle is associated with the highest PTSD rates. This may be compounded by the impact of life-changing severe injuries such as traumatic brain injury (TBI), traumatic amputations, spinal cord injuries, and blindness.

A common denominator for many returnees is the sustained hypervigilance and persistent anxiety about potential threats at any time or place, which often leads to difficulty transitioning from the theater back to civilian life. Clinicians faced with patients who have had a difficult re-entry to the home front need to be aware of this reintegration process. They must consider the likelihood that postdeployment difficulty for a particular patient may be part of an otherwise normal readjustment trajectory or could be manifestations of a clinically significant problem such as PTSD.

Assessment and Diagnosis

A routine assessment should begin with a few questions about the patient's exposure to war zone trauma and other traumatic experiences. Once the presence of traumatic exposure is established, the following step is to screen for PTSD. The National Center for PTSD has developed a four-item (yes/no) screen with excellent psychometric properties (Primary Care PTSD Screen [PC-PTSD]).

Designed for use by primary care practitioners, the PC-PTSD consists of four questions, one from each PTSD symptom cluster, asked in reference to the past month: (1) Have you had nightmares or thought about the trauma when you did not wish to? (re-experiencing), (2) Have you tried to avoid thoughts and situations that remind you about the event? (avoidance), (3) Are you easily startled or frequently on guard? (hyperarousal), and (4) Have you felt numb or detached from other people or things? (numbing). A patient that endorses three of the four items should receive a more thorough assessment for PTSD. Recent research with the PC-PTSD shows that it correctly positively identifies those that have PTSD 78% of the time (sensitivity) and correctly negatively identifies those that do not have PTSD 87% of the time (specificity).

Given sufficient evidence from screening that the patient may have PTSD, a more formal assessment is warranted. This can be accomplished through a comprehensive diagnostic interview in which each of the 17 PTSD symptoms is assessed, by use of a structured clinical interview such as the Clinician Administered PTSD Scale (CAPS) or using well-validated self-report questionnaires such as the PTSD Checklist (PCL), the PTSD Diagnostic Scale (PDS), the Davidson Self-Rating PTSD Scale, or others.

There are a few things that clinicians need to keep in mind when dealing with veteran patients. Clinicians should be aware that service members often fear being stigmatized if they disclose any sign of psychiatric problem, including PTSD. Furthermore, those that are most symptomatic are more likely to feel stigmatized and therefore less likely to seek treatment. It is also necessary to recognize that although social support is a powerful protective factor against developing PTSD, it is much less available to National Guard and military reservists and their families than to full-time military members. Whereas the latter may inhabit military bases where many military families live close together, undergoing the same deployment stressors, reservists, and especially their families, are neither embedded in a full-time military culture nor surrounded by others who share their plight. This may explain why reserve troops in the Persian Gulf War had higher prevalence of PTSD and depression than active duty personnel.

Treatments

Effective evidence-based psychotherapeutic and pharmacological treatments for PTSD are available. Detailed discussion of the empirical treatment literature can be found in clinical practice guidelines.

Psychotherapeutic Treatments

Clinical guidelines published to date designate cognitive-behavioral therapy (CBT), which includes cognitive therapy, prolonged exposure, and cognitive processing therapy, as the treatment of choice for PTSD. CBT techniques are designed to alter the conditioned fear and cognitive distortions associated with PTSD. Prolonged exposure essentially involves clinicians asking patients to construct narratives about intolerable traumatic events they can recall.

With repeated, therapist-guided exposure to these traumatic memories (imaginal or *in vivo*), patients experience progressive reduction in distress and thereby achieve clinical remission. Cognitive therapy and cognitive processing therapy focus on changing the maladaptive automatic thoughts associated with PTSD. Cognitive restructuring is a technique through which the therapist challenges distorted beliefs (perceiving the world as dangerous, seeing oneself as powerless or inadequate, or feeling guilty for outcomes that were outside of one's control), thereby enabling patients to overcome intolerable trauma-related emotions. Exposure and cognitive therapies have performed very well and with equal efficacy.

Eye movement desensitization and reprocessing (EMDR) is a technique that involves patients being instructed to imagine a painful traumatic memory and associated negative cognitions while visually focusing on the rapid movement of the clinician's finger. Although there are now many studies showing that such eye movements are not needed for EMDR to work, evidence-based practice guidelines indicate that EMDR is an effective treatment despite the lack of a defined mechanism of action.

In practice, both exposure and cognitive therapies are considered first-line treatments for PTSD. In exposure therapy, a clear memory of the traumatic event must be accessible. For cognitive therapy, the focus is primarily on the automatic, erroneous cognitions and behaviors that are associated with traumatic memories. In practice, exposure therapy has successfully addressed erroneous cognitions, and cognitive therapy has successfully addressed fear-based avoidant behavior. Therefore, at this time, the choice of CBT therapy for PTSD is up to the clinician. Given the scarcity of trained CBT therapists, neither option of evidence-based treatment may be available.

Pharmacological Treatments

A number of medications have been tested as treatment for PTSD patients. Selective serotonin reuptake inhibitors (SSRIs) are currently the treatment of choice, and the U.S. Food and Drug Administration has approved two, sertraline and paroxetine, for the treatment of PTSD. What has been especially exciting about the SSRI results is that these drugs appear to be broad spectrum agents that work to reduce all the symptom clusters of PTSD. Successful randomized clinical trials have also been conducted with the SSRI fluoxetine, the alpha-1 adrenergic antagonist, prazosin, tricyclic antidepressants, monoamine oxidase inhibitors, and venlafaxine. Because the psychobiology of PTSD is complex, other medications such as antiadrenergic agents, anticonvulsants, and a variety of agents currently under investigation may prove as effective as SSRIs. Finally, randomized trials of benzodiazepines have yielded negative results, so this class of medication cannot be recommended for PTSD treatment. When choosing medications and other elements of treatment, keep in mind that PTSD is highly comorbid with other psychiatric problems.

Summary

Many important questions about treatment have yet to be tested systematically. These include how to select treatment, define realistic goals, combine various treatments, and approach complex comorbid conditions; how long to stick with a specific treatment; and when to acknowledge clinical failure. As new psychotherapeutic and pharmacological treatments are developed and empirically tested, these essential questions should also be addressed. In the meantime, it is essential that clinicians select evidence-based treatments whenever possible.

See Also the Following Articles

Acute Stress Disorder and Posttraumatic Stress Disorder; Cognitive Behavioral Therapy; Vietnam Veterans, Postwar Experiences and Health Outcomes; Selective Serotonin Reuptake Inhibitors (SSRIs).

Further Reading

American Psychiatric Association (2000). Practice guidelines for the treatment of acute stress and posttraumatic stress disorder. *American Journal of Psychiatry* **161**, 1–31.

Asnis, G. M., Kohn, S. R., Henderson, M., et al. (2004). SSRIs versus non-SSRIs in post-traumatic stress disorder: an update with recommendations. *Drugs* **64**, 383–404.

Brewin, C. R., Andrews, B. and Valentine, J. D. (2000). Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. *Journal of Consulting and Clinical Psychology* **68**, 748–766.

Foa, E. B., Keane, T. M. and Friedman, M. J. (2000). *Effective treatments for PTSD: practice guidelines from the International Society of Traumatic Stress Studies*. New York: Guilford Press.

Friedman, M. J. (2002). Future pharmacotherapy for PTSD: prevention and treatment. *Psychiatric Clinics of North America* **25**, 427–441.

Friedman, M. J. (2005). Veterans' mental health in the wake of war. *New England Journal of Medicine* **352**, 1287–1290.

Friedman, M. J. (2005). Acknowledging the psychiatric cost of war. *New England Journal of Medicine* **351**, 75–77.

Hoge, C. W., Castro, C. A., Messer, S. C., et al. (2004). Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *New England Journal of Medicine* **351**, 13–22.

Kulka, R. A., Schlenger, W. E., Fairbank, J. A., et al. (1990). *Trauma and the Vietnam War generation: report of findings from the National Vietnam Veterans Readjustment Study*. New York: Brunner/Mazel.

Murdoch, M., Polusny, M. A., Hodges, J., et al. (2004). Prevalence of in-service and post-service sexual assault among combat and noncombat veterans applying for Department of Veterans Affairs posttraumatic stress disorder disability benefits. *Military Medicine* **169**, 392–395.

National Collaborating Centre for Mental Health (2005). *Post-traumatic stress disorder: the management of PTSD in adults and children in primary and secondary care*. London: Gaskell and the British Psychological Society.

Prins, A., Ouimette, P. C., Kimerling, R., et al. (2004). The primary care PTSD Screen (PC-PTSD): development and operating characteristics. *Primary Care Psychiatry* **9**, 9–14.

Resick, P., Nishith, P., Weaver, T., et al. (2002). A comparison of cognitive processing therapy with prolonged exposure and a waiting condition for the treatment of chronic posttraumatic stress disorder in female rape victims. *Journal of Consulting and Clinical Psychology* **70**, 867–879.

Shapiro, F. and Maxfield, L. (2002). Eye movement desensitization and reprocessing (EMDR): information processing in the treatment of trauma. *Journal of Clinical Psychology* **58**, 933–946.

VA-DoD Clinical Practice Guideline Working Group, Veterans Health Administration, Department of Veterans Affairs and Health Affairs, Department of Defense: Management of post-traumatic stress. Washington, D.C: Office of Quality and Performance publication 10Q-CPG/PTSD-04 2003.

Wilson, J. P. and Keane, T. M. (2004). *Assessing psychological trauma and PTSD* (2nd edn.). New York: Guilford Press.

Weapons of Mass Destruction *See:* Hiroshima Bombing, Stress Effects of.

Women and Stress *See:* Sex Differences in Human Stress Response; Stress Induced Anovulation.

Work–Family Balance

J G Grzywacz
Wake Forest University School of Medicine, Winston-Salem, NC, USA
A B Butler
University of Northern Iowa, Cedar Falls, IA, USA

© 2007 Elsevier Inc. All rights reserved.

Defining Work–Family Balance
Prevalence and Emergence of Interest in Work–Family Balance
Work–Family Balance and the Stress Process
Promoting Work–Family Balance
Gaps and Needed Research

Glossary

Family	Responsibilities and activities that arise from shared goals, values, and long-term commitments among two or more people who are related by common ancestry, adoption, marriage, and other legal or socially recognized unions.
Work	Responsibilities and activities inherent in an individual's trade, profession, vocation, or other means of livelihood.
Work–family conflict	The extent to which an individual's work interferes with his or her family or an individual's family interferes with his or her work.

Work–family facilitation	The extent to which an individual's work benefits his or her family or an individual's family benefits his or her work.
Work–family interface	Areas of physical, psychological, social, or temporal overlap in which an individual's work and family affect each other.

Defining Work–Family Balance

Work–family balance remains largely undefined despite its frequent use in both the popular and scholarly vernacular. One influential definition characterizes work–family balance in terms of equal engagement in both work and family responsibilities. This definition of work–family balance is influential because it is clearly grounded in theory and is consistent with the balance metaphor. Despite these strengths, however, the definition overlooks several important complexities such as the multiple forms or ways in which individuals can engage in their work and family lives, and whether or not different forms of engagement are functionally equivalent. Recognizing these complexities, more recent definitions of work–family balance continue to highlight notions of equality, but they also emphasize comparable levels of satisfaction within each domain and the relative absence of conflict between the domains. Work–family balance, therefore, generally refers to a work and family life that is well integrated and a situation in which the individual is actively engaged in both domains.

Prevalence and Emergence of Interest in Work–Family Balance

Difficulty maintaining work–family balance is reported by the majority of working adults. Estimates from the U.S. General Social Survey suggest that fewer than 40% of adults believe that they are successfully balancing their work and family lives. The widespread difficulty in achieving work–family balance, and subsequent scholarly and popular interest, results from significant socioeconomic shifts in the constitution of the labor force as well as in the nature of work itself. In the final quarter of the twentieth century, women's labor force participation increased dramatically, particularly among women with children. The proportion of children being raised in a single-parent household increased, as did the rate of employment among single parents. Concomitantly, substantial job growth in the service-producing sector coupled with increased global competition created new demands on the labor force, such as increased use of contingent labor and other flexible staffing models, as well as greater need for work that occurs outside the bounds of traditional business hours.

The shifts in work and family in the later portion of the twentieth century made navigating daily work and family life more complex for many adults. Growth in women's labor force participation challenged the gender-based division of labor and brought basic issues related to household management and childcare to the forefront. Securing high-quality and affordable childcare is an ongoing challenge, as is care for sick children and elder care. The transition to a service-based economy was accompanied by a fall in real wages, particularly among lower and working-class families. Globalization and the increased reliance on contract and contingent employment arrangements elevated concerns over job insecurity and raised new questions about the nature of the employer–employee relationship. Technology and the 24/7 economy has redefined when, where, and how work is performed for a substantially larger portion of workers. The increased complexity of daily work and family life for many adults has fueled widespread interest in the causes and consequences of work–family balance.

Work–Family Balance and the Stress Process

Work–family balance has been applied and studied at two major points in the stress process. Work–family balance or, more typically, imbalance, is frequently conceptualized as a primary stressor. The basic theoretical rationale for viewing imbalance between work and family as a primary stressor is that the social roles of employee/worker, spouse, and parent have substantial social and cultural value and, therefore, play a central role in adults' identity and sense of self. Indeed, some researchers suggest that successful integration of work and family life is a central task of adult development. If work and family are out of balance, there is substantial social and personal pressure placed on individuals to better manage their responsibilities.

Multiple lines of research support the view of work–family imbalance as a primary stressor. Adults seek to maintain work–family balance through a variety of strategies such as changing jobs, reducing hours, renegotiating household responsibilities, cutting back on personal leisure pursuits, and modifying standards or expectations for their work and family life. Individuals who do not have a well-balanced work and family life – typically operationalized in terms of high levels of work–family conflict – have less life satisfaction as well as elevated levels of depression and somatic complaints, conditions frequently implicated in the stress response. Corroborating this research is more recent evidence linking indicators of work–family balance to common physiological stress responses such as elevated blood pressure, heart rate,

and cortisol. Social support from co-workers, supervisors, or family members has been found to buffer the effects of poor work–family balance on stress-related outcomes such as levels of distress and domains of life satisfaction. In summary, there is a wide evidence base suggesting that work–family balance behaves like a primary stressor.

Other research situates work–family imbalance as a secondary stressor, or a downstream consequence of other circumstances or events. This research exists in two largely separate streams of inquiry. In the first stream, work–family balance is studied as a consequence resulting from the relative availability of work- and family-related resources to meet the demands of work and family life. This research focuses on work-related demands such as the time and timing of when work is performed, features of the work such as the amount of authority and variety workers are given, the availability of family-friendly policies such as flexible work schedules or on-site childcare centers, and the supportiveness of co-workers and supervisors. On the family side, research tends to focus on the number and ages of children, the employment status of adults in the family (e.g., dual-earner versus single-earner), qualities of intrafamily relations such as levels of conflict among partners, and the division of household labor. In the second stream of research, work–family balance is seen as a mechanism through which broader social forces exert influence on individual outcomes. For example, some investigators argue that observed health disparities, particularly in child health and development, result (at least in part) from the social allocation of demands and resources necessary for successfully balancing work and family. Racial and ethnic minorities as well as individuals with a low education, for example, are overrepresented in jobs requiring nonstandard schedules that provide few (if any) formal programs that assist workers in balancing their work and family responsibilities. Despite these arguments and the fact that the current focus on work–family balance is a consequence of profound sociostructural changes in work and family, the literature is dominated by views of balance as a personal challenge confronted by individuals.

Promoting Work–Family Balance

Given the stress-related consequences of work–family imbalance, both individuals and organizations may be motivated to take action to restore or maintain balance. Individuals who set goals and employ adaptive strategies to reach those goals are less likely to experience conflict between work and family. An active coping style and social support are also associated with greater work–family balance. Organizational interventions to improve balance typically focus on offering family-friendly benefits designed to facilitate management of conflicting role demands. Such benefits are broad and varied, including flexible scheduling, on-site daycare, and eldercare assistance. These benefits tend to be underutilized by workers, and evidence regarding their efficacy for improving balance is equivocal. Family-supportive organizational cultures, of which family-friendly benefits are but one feature, are also widely believed to be essential to promoting work–family balance.

Gaps and Needed Research

Understanding of work–family balance is undermined by several conceptual and methodological gaps in the literature. Conceptually, although most researchers agree that work–family balance represents some vague notion of equality and harmony in an individual's work and family life, it is not clear what work–family balance is or whether the balance metaphor is appropriate. Moreover, there is very little theoretical work underlying work–family balance research. Methodologically, measurement of work–family balance remains encumbered. Most studies rely on single-item measures (e.g., How successful do you feel at balancing your paid work and your family life?). There is currently a division in the literature regarding how to measure work–family balance: some scholars advocate the use of individuals' overall appraisals of balance, whereas others suggest that work–family balance is a latent construct best measured indirectly using indicators of work–family conflict and work–family facilitation. Future research needs to systematically evaluate the strengths and weaknesses of each approach and develop recommendations. Although there are notable exceptions, work–family balance research is also overly reliant on cross-sectional cohort designs, convenience samples, and self-report measures. Research is needed that clarifies the conceptual meaning of work–family balance, develops valid and innovative measures of work–family balance, and studies the phenomenon using multiple methods.

See Also the Following Articles

Burnout; Control and Stress; Environmental Stress, Effects on Human Performance; Familial Patterns of Stress; Fatigue and Stress; Gender and Stress; Social Status and Stress; Social Support; Workplace Stress.

Further Reading

Clark, S. C. (2000). Work/family border theory: a new theory of work/family balance. *Human Relations* **53**, 747–770.

Eckenrode, J. and Gore, S. (eds.) (1990). *Stress between work and family.* New York: Plenum.
Frone, M. R. (2003). Work-family balance. In: Quick, J. C. & Tetrick, L. E. (eds.) *Handbook of occupational health psychology*, pp. 143–162. Washington, D.C.: American Psychological Association.
Greenhaus, J. H., Collins, K. M. and Shaw, J. D. (2003). The relation between work-family balance and quality of life. *Journal of Vocational Behavior* **63**, 510–531.
Heymann, J. (2000). *The widening gap: why America's working families are in jeopardy, and what can be done about it.* New York: Basic Books.
Marks, S. R. and MacDermid, S. M. (1996). Multiple roles and the self: a theory of role balance. *Journal of Marriage and the Family* **58**, 417–432.
Voydanoff, P. (2005). Toward a conceptualization of perceived work-family fit and balance: a demands and resources approach. *Journal of Marriage and Family* **67**, 822–836.

Workplace Stress

U Lundberg
Stockholm University, Stockholm, Sweden

© 2007 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by U Lundberg, volume 3, pp 684–692, © 2000, Elsevier Inc.

Introduction
Impact of the Work Environment
Models of Work Stress
Type A Behavior
Gender Differences in Response to Work Stress
Psychobiological Mechanisms Linking Work Stress to Health Problems
Concluding Remarks

Glossary

Allostatic load	The cost of adaptation (allostasis), reflecting wear and tear on the body systems and processes that increases the risk for disease.
Cardiovascular illness	Circulatory disturbances such as hypertension, myocardial infarction, and stroke.
Demand–control model	Defines work strain as a combination of high workload and low control. The dimension of social support has been added to this model.
Effort–reward imbalance model	Stress and health risks are created by an imbalance between high effort and lack of adequate rewards (money, status, support).
Musculoskeletal disorders	Degenerative processes, damage, inflammation, and pain in muscles and joints in the body (neck, shoulder, arms, lower back, etc.).
Sympathetic-adrenal-medullary system and the hypothalamic-pituitary-adrenocortical axis	These form the two major neuroendocrine stress systems in the body.

Introduction

Work has a very central role in most people's lives. In industrialized countries, there have been quite dramatic changes in the conditions at work during the past decades, caused by economic, social, and technical development. Today, as a consequence, people at work are exposed to high quantitative as well as high qualitative demands and to hard competition, caused by a global economy, multinational companies, lean production, downsizing, and increased demands for efficiency. Fewer employed people are expected to produce more, and experiences of stress due to overstimulation have become a serious problem. At the same time, a considerable part of the population is unemployed or only temporarily employed, because no jobs are available or because they have an inadequate or insufficient education. These people experience stress mainly due to lack of economic resources and understimulation. This polarization of the labor market is likely to have contributed to the social gradient in health and well-being, which has increased in recent years. This article summarizes present knowledge regarding how the modern work environment may contribute to stress and negative health outcomes.

The Role of Work

Most adults spend a large part of their daily life at work. As a consequence, conditions at the workplace

are likely to be of considerable importance for people's mental as well as physical well-being. Work has a central role in people's lives, and a stimulating job may contribute to a more meaningful life, a higher sense of self-esteem, social ties, and economic independence – conditions that characterize positive human health. Work also creates a time structure in daily life and is a major determinant of the individual's socioeconomic status. Negative psychosocial conditions at work, such as a low job satisfaction and lack of influence and control, are known to be associated with various health risks.

Changing Work Conditions

During the past century, the work situation in industrialized countries has undergone profound changes. For women in particular, these changes have been quite dramatic. At the beginning of the twentieth century, women were often involved in heavy physical work (e.g., working in the fields as a farmer's wife). As urbanization increased in the middle of this century, a married woman was generally expected to assume the role of homemaker, with the primary duty of taking care of her children, husband, and home or household. In the last part of this century, an increasing number of women entered the labor market, and in the Scandinavian countries (as well as in many big cities all over the world) the labor force today comprises practically the same number of women as men. However, despite women's greater involvement in paid work and their importance for the family economy, their responsibility for unpaid work at home does not seem to have been reduced accordingly. In keeping with this, employed women often report stress from work overload and role conflicts.

In the modern work environment, physical hazards and demands have been reduced, whereas psychosocial stress, caused by a very high work pace, competition and efficiency, and successive readjustment to organizational changes, has increased. This can be seen as part of a general and accelerating trend in modern society toward a more rapid pace of life, more frequent introduction of new technology, and continuous demands for coping with new conditions. Major improvements in physical conditions at work during the past decades have thus contributed to a switch in health focus, specifically to the role of psychosocial factors such as stress.

Developments in communication technology (faxes, e-mail, Internet, mobile phones, etc.) have enabled greater flexibility and mobility (e.g., teleworking) and have removed traditional boundaries between different roles in life (work, family, leisure). In addition, temporary employment and work on short-term projects have become more common. These new trends are likely to involve the potential for beneficial effects in terms of greater variation and flexibility, but also an increased risk of stress due to work overload, role conflicts, and lack of time for rest and recovery. The individual, rather than the company, has to set the boundaries between work and his or her other roles in life.

Impact of the Work Environment

Conditions at the workplace may influence the individual in different ways. Physical conditions, such as noise, light, cold, heat, and so on are sensorily mediated by the individual and directly perceived as, for example, harmful or pleasant. In contrast, the individual cannot directly perceive exposure to many other physical conditions, such as radiation, certain chemicals and heavy metals, air and water pollution, and radioactive fallout, as harmful. The health risks associated with such work conditions are mainly based on information obtained from education, mass media, personal communication, or earlier experiences. In such cases, it is not often clear to what extent the stress and health problems associated with these occupational hazards are caused by the actual exposure or by the individual's own fear and anxiety from having been exposed.

Health problems can be caused by the effects of stress on various bodily systems, functions, and organs, such as blood pressure, secretion of stress hormones, immune functions, metabolism, the heart, the brain, and the gastrointestinal system. Health risks can also be caused by the effects of stress on various health-related behaviors, such as cigarette smoking, alcohol and drug abuse, poor eating habits, and lack of physical activity. Additional health risks at work are associated with violence or accidents caused by the greater risk of making mistakes under stress, or being reluctant (or neglecting) to use protective means (e.g., seat belts, helmets, hearing protection) when under time pressure.

Associations between work stress and cardiovascular illness (e.g., hypertension, myocardial infarction) have been reported repeatedly for a long time. However, there is also increasing evidence for strong associations between psychosocial work conditions and a great number of other health problems. A major work-related health problem today is that of musculoskeletal disorders. Neck, shoulder, and lower back pain are the most common reasons for absenteeism from work and for early retirement in Europe and North America, and also affect younger workers. In addition to the pain and suffering caused by these disorders, the economic costs are enormous.

Epidemiological studies have demonstrated dramatic and very consistent relationships between socioeconomic status and health. Social status is usually defined by type of occupation and educational level, and higher status is consistently linked to better health. The social gradient in health can only partly be explained by differential exposure to physical hazards at work, poor material conditions (food, housing), genetic factors, or health-related behaviors. Other factors must also be involved, and there is strong evidence that psychosocial factors at work and income inequality are important for the social differences in health.

Models of Work Stress

Several models have been proposed to explain the causes of work-related stress. Frankenhaeuser and colleagues and other investigators have described a model in which stress is defined in terms of the imbalance between the perceived demands from the environment and the individual's perceived resources to meet those demands. This imbalance can be caused by quantitative overload (a very high pace of work, too much work to do, etc.) or qualitative overload (too much responsibility, problems too complex to solve, conflicts, etc.). However, an interesting feature of this model is that it also proposes that stress may be due to an imbalance caused by understimulation, for example, by work tasks that are too simple, that do not allow individuals to use their education, skills, and experiences, and that do not allow them to learn new skills and develop their abilities. This situation can be found in monotonous and repetitive work, such as traditional assembly line work and in data entry work at video display units, and among people who are unemployed. It is also common among highly educated immigrants, who often have to accept very simple jobs due to insufficient knowledge of the new language or discrimination, which means that their training and experience cannot be used.

Another well-known model describing work stress or strain is the demand–control model, proposed by Karasek and Theorell and developed and expanded by others. According to this model, the combination of high demands and lack of control and influence (low job discretion) over the work situation causes high work strain. High demands combined with a high degree of control, which characterizes many high-status jobs, are described as an active work situation and are not associated with enhanced health risks.

The demand–control model has been tested in numerous studies that, in general, show that occupations characterized by high job strain, for instance, a waiter or an assembly worker, are associated with elevated health risks compared with low-strain jobs, such as being a forester or a scientist. Although most studies are cross-sectional, thus excluding the possibility of making causal inferences, the few prospective studies that have been performed report similar findings. In the late 1980s, a third dimension – social support – was added to the demand–control model. High job strain combined with low social support at work contributes to even more elevated health risks.

The control dimension describes a wide range of factors typical of a positive and healthy work situation. It involves the individual's possibilities for developing and using his or her skills, knowledge, education, and experience to be able to work out new methods and ways of doing the job, learning new skills, taking responsibility, being independent, and experiencing variation at work. Low scores on this dimension have consistently been linked to various negative health outcomes, whereas the results regarding the dimension demands are less conclusive. A possible explanation for the inconsistent findings regarding high demands could be that this dimension does not distinguish between physical and mental demands or between quantitative (e.g., high work pace) and qualitative demands (e.g., too much responsibility). For example, individuals in high-ranking positions often report both high demands and good health.

Social support is generally considered to protect against stress at work or to serve as a buffer against health risks under stressful conditions. It is defined by factors such as having close social ties and someone to share emotional experiences with, good collaboration with colleagues and superiors, the possibility of getting help when needed, and adequate feedback about one's effort and performance. The number of social contacts is less important than the quality of these relationships. In a literature review, low social support was consistently found to be associated with higher morbidity and mortality.

Siegrist proposed a new model for stress at work called the effort–reward imbalance model. According to this model, lack of adequate reward in response to the individual's achievement efforts is considered to contribute to high stress levels and elevated health risks. Reward could be obtained in terms of economic benefits such as a higher income, but also in terms of appreciation and adequate support from colleagues and superiors, or by obtaining better career chances and a higher social rank at the job. A personality characteristic, overachievement, is also an important part of this model; individuals high in overachievement are at particular risk for health problems. In a

recent prospective study, this model was found to significantly predict elevated risk of myocardial infarction.

Type A Behavior

A stress- and work-related behavior pattern termed type A behavior has attracted a considerable amount of attention over the past decades. This attention was based, to a large extent, on findings from the prospective Western Collaborative Group Study, in which about 3000 middle-aged employed men were classified either as type A individuals or as their more relaxed type B counterparts on the basis of a standardized interview technique (the structured interview). At the 8.5-year follow-up, it was found that the prevalence of myocardial infarction was about twice as high among type A as among type B individuals, after controlling for traditional risk factors (blood pressure, blood lipids, cigarette smoking, etc.). This study initiated a great number of investigations and experiments aimed at revealing factors that contribute to the development of this behavior pattern and the psychophysiological mechanisms that could create a link to myocardial infarction. Work-related stress was found to play an important role in this context.

Type A behavior is defined in terms of an extreme sense of time urgency, impatience, competitiveness, and aggression/hostility. Work conditions in industrialized countries, emphasizing the importance of efficiency, productivity, a high work pace, and competitiveness, are considered to contribute to the development of this behavior pattern. In addition, individuals classified as type A have been found to be particularly responsive to challenges at work (obstacles to keeping up the pace of work, lack of control, competition, harassment, etc.) in terms of blood pressure, heart rate, blood lipids, and catecholamine output. Intensive, frequent, and/or sustained activation ofthese physiological stress responses contributes to the atherosclerotic process and to blood clotting and are, thus, likely to form a pathway between type A behavior and the elevated risk of myocardial infarction. Results for women are in general less consistent than they are for men, which could be explained by the fact that the structured interview was developed using male subjects.

The high work pace and competitiveness associated with the type A behavior pattern are usually reinforced at work by giving the individual a higher income, appreciation, and a successful occupational career (higher rank). Consequently, efforts to modify this behavior pattern in healthy individuals have not been particularly successful. However, after a myocardial infarction individuals tend to be more motivated to change their lifestyle, and, indeed, a prospective intervention study has shown a reduced risk of reinfarction among individuals who were able to modify their type A behavior.

Interest in the global type A behavior has diminished in favor of one specific component of this behavior pattern – the aggression-hostility component, which seems to be the most toxic factor in terms of myocardial risk. The important role played by hostility in cardiovascular disease was first demonstrated in a longitudinal study by Barefoot and colleagues, which found that medical students with high scores on the Cook-Medley hostility scale of the Minnesota Multiphasic Personality Inventory (MMPI) had a sixfold increase in mortality when followed up 25 years later, mainly due to coronary heart disease (CHD). Subsequent studies have consistently supported this finding by showing that individuals who are high in hostility have a significantly increased risk of myocardial infarction. As medication to prevent and treat cardiovascular disease has improved considerably during the past decades and the incidence rate of myocardial infarction has declined steadily since 1980, interest in performing large prospective studies on type A behavior and CHD has diminished accordingly.

Gender Differences in Response to Work Stress

As the number of women in the labor force approaches that of men but the traditional female responsibility for home and family mainly remains the same, stress from work overload and role conflicts has become an increasing problem for many women. Although women generally live longer than men, women tend to report more health problems than men. Women's longevity is associated with behavioral factors (e.g., men use more drugs and alcohol, men commit suicide more often, men expose themselves to greater risks at work and in traffic), environmental factors (men perform more dangerous jobs, are more violent, and engage in criminal behavior more often) and biological factors (women contract fatal cardiovascular disorders such as myocardial infarction and stroke about 10 years later in life than men, mainly due to the protective effects of female sex hormones before menopause). However, musculoskeletal and gastrointestinal disorders, headache, chronic fatigue, and psychiatric disorders are reported more frequently by women than by men.

Women's employment, even in low-status jobs, seems to be associated with better health compared to the role of a homemaker. To some extent this could

be explained by the fact that it is easier for women in good health to get and keep a job, compared with women with poorer health. However, a paid job is also likely to contribute to better health and well-being by increasing the possibilities for social relations, higher self-esteem, and economic independence. Provided that men and women are matched for education and occupational level, physiological stress responses at work seem to be quite similar. However, consistent gender differences have been found off the job. Frankenhaeuser and colleagues compared stress responses during and after work in male and female white collar workers, matched for age and occupational level, and found that men's physiological arousal rapidly returned to baseline after work, whereas women's stress levels remained high several hours after work. These findings have been replicated in recent studies of men and women in high-ranking positions. A likely explanation for these gender differences is that women were unable to unwind and relax due to their greater responsibility for unpaid work at home (e.g., household chores, child care).

Several studies show that women's work stress is reflected not only at work but also in terms of their physiological arousal in non-work conditions. Men seem to respond more specifically to the acute stress exposure at work, and their work stress does not seem to spill over into non-work conditions.

In a prospective study of approximately 600 000 men and 400 000 women in Sweden, Alfredsson and colleagues found that overtime at work (10 h or more per week) was associated with a significantly elevated risk of myocardial infarction during a 1-year follow-up in women but not in men. Men's overtime was even associated with a significantly reduced risk of infarction.

One possible explanation for this gender difference, which is supported by psychophysiological data in other studies, is that women's CHD risk is related to elevated and sustained physiological stress levels caused by work overload and their more pronounced role conflicts, when they are trying to combine overtime at work with responsibility for various unpaid duties at home. Another possible explanation is that women, more often than men, are employed in low-status jobs, in which high demands are combined with little influence over the work situation, including overtime. Such high-strain jobs, according to the demand–control model, are known to increase the risk of CHD in both men and women.

A paid job as well as the unpaid work of household chores and child care is likely to contribute to the greater total workload of full-time employed women, compared with men. This could be relevant for women's greater health problems, but it also reduces their chances of having a professional career on the same terms as men. A study of about 1000 male and 1000 female full-time employed white collar workers, matched for age and type of occupation, confirmed that a greater proportion of women compared to men reported having the main responsibility for most unpaid duties at home. In terms of number of working hours in paid and unpaid productive activity, the total workload of men and women seems to be about the same in families having no children at home, whereas the workload increases more rapidly for women than for men depending on the number of children at home.

Psychobiological Mechanisms Linking Work Stress to Health Problems

Allostatic Load

A new stress model, the allostatic load model, refers to the ability to achieve stability through change. A normal and economic response to stress means activation of physiological systems in order to cope with the stressor, followed by a shut-off of the allostatic response as soon as the stress is terminated, whereas over- or underactivity of the allostatic systems may add to the wear and tear of the organism. Examples are frequent and intensive activation of physiological systems (cardiovascular, hormonal, metabolic, immune function) without enough time for rest and recovery or an inability to shut off the stress response after the stress exposure, which may cause overactivity and exhaustion of the systems. Overexposure to stress hormones increases the risk of cardiovascular illness and immune deficiency, as well as cognitive impairment. Lack of adequate response in one system, due to exhaustion, may cause compensatory overactivation of other systems.

The allostatic load associated with blue collar jobs is usually higher than that associated with white collar jobs. Manual jobs involve mental as well as physical stress and are associated with more pronounced physiological responses. In addition, repetitive work associated with many blue collar jobs seems to contribute to sustained levels of stress after work, which further adds to the allostatic load. In women, the unpaid work responsibilities of household chores and child care may further contribute to a lack of opportunity for unwinding and to sustained stress levels in non-work situations.

Health problems are consistently more common among blue collar workers with low education, responsibility for monotonous and repetitive tasks, and low income than among white collar workers with more stimulation and variation at work.

Work Stress and Activity of the Hypothalamic-Pituitary-Adrenocortical Axis

Negative psychosocial and socioeconomic factors are associated not only with increased morbidity and mortality but also with elevated activity of the hypothalamic-pituitary-adrenocortical (HPA) system and increased cortisol secretion. Additionally, a very high workload, such as regularly working more than 10 h of overtime per week, has been associated with markedly elevated cortisol levels, at least in women. Sustained activity of the HPA system is related to a series of endocrine and metabolic effects, causing, among other things, increased storage of fat in the abdominal region. The amount of central fat accumulation indicates long-term overactivity of the HPA system and can be measured as the relationship between the waist and the hip circumference, the waist–hip ratio (WHR). An association has been demonstrated between high WHR and a series of negative socioeconomic factors, such as low education and income, poor housing and/or living alone, psychiatric problems and use of antidepressive drugs, and high absenteeism from work. The relationship between low socioeconomic status and high WHR is enhanced by the fact that health-related behaviors, such as cigarette smoking, alcohol and drug abuse, and lack of exercise, also contribute to increasing the activity of the HPA system. High WHR has been shown to increase the risk of diabetes and cardiovascular disease, and, through its anti-inflammatory effects, cortisol also has an important role in the functioning of the immune system. It has been concluded that chronic or long-term stress is associated with compromised immune functions and, thus, increased risk for infections and slower healing of wounds. Elevated cortisol levels have recently also been related to cognitive functions, such as explicit memory deficiency and low hippocampal volume.

Thus, in summary, the pattern of results suggests that individuals exposed to negative psychosocial work conditions (low income, low education, risk of unemployment, etc.) are more likely to have elevated cortisol secretion, more central fat accumulation, and greater health risks.

Work Stress and Musculoskeletal Disorders

Musculoskeletal disorders increased dramatically during the 1980s and have become one of the most important work-related health problems in Europe and North America today. Research indicates that psychosocial work stress is an important factor contributing to this development.

The association between such factors as bad ergonomic conditions at work, heavy lifting, and physically monotonous or repetitive work and an increase in shoulder, neck, and lower back pain problems has been well documented. However, considerable ergonomic improvements of the work environment have not resulted in a lower incidence of musculoskeletal disorders, and musculoskeletal disorders are frequent not only in physically demanding jobs but also in light physical work. An increasing number of studies report an association between psychosocial factors at the workplace and neck, shoulder, and lower back pain problems. Conditions typical of many low-status jobs, such as low work satisfaction, time pressure, lack of influence over one's work, and constant involvement in repetitive tasks of short duration, often characterize jobs associated with a high risk for muscular problems.

Musculoskeletal disorders differ from many other major health problems, such as cardiovascular disease and cancer, in that symptoms often appear very early in life and after a relatively short exposure to adverse environmental conditions. In repetitive work, pain syndromes are often reported after only 6–12 months on the job, and it has been documented that even very low levels of muscle activation (less than 1% of maximal voluntary contraction) may contribute to the development of chronic pain syndromes.

So far it has been difficult to explain the high incidence of musculoskeletal disorders in light physical work. However, it has been suggested that sustained low-level muscle activity, induced by physical and/or psychological demands, may initiate a pathogenetic mechanism resulting in muscle pain and thus form a link to musculoskeletal disorders. In a recent experimental study, it was demonstrated that the smallest functional units of the muscle, the motor units, can be activated by physical as well as mental stress. Psychological and psychosocial factors on and off the job, muscle pain, and illness behavior may also form a vicious circle with increasing sensitization and pain symptoms. Eventually, the individual will become chronically ill.

An important feature of psychological stress, which could be relevant to its role in the etiology of muscle pain, is that it is more lasting or chronic than physical demands at work. Stressful conditions such as low job satisfaction, low income, low status, an irritating boss, and fear of losing one's job influence the individual more or less continuously, whereas physical demands terminate at the end of the workday and during breaks at work.

Support for the assumption that work stress contributes to muscle tension and pain has been obtained inseveral studies. For example, in an interesting prospective study of female packing workers, it was found that women who contracted clinically diagnosed

trapezius myalgia within the first year of employment had significantly higher muscle activity during breaks at work but not during actual work and, in addition, had fewer periods of very low muscular electrical activity (electromyographic [EMG] gaps). These findings indicate that, in the modern work environment, lack of relaxation during breaks at work and after work may be an even more important health problem than the actual level of stress or the intensity of muscle activation during work.

The prevalence of musculoskeletal disorders is markedly higher among women than among men. One would generally expect that this gender difference could be explained by women's lower physical capacity compared with men. However, because several studies show that the physical characteristics of the individual do not seem to predict future musculoskeletal disorders and that the physical load in white collar jobs, where the gender difference is most pronounced, is usually quite low, a more likely explanation is that women are often employed in stressful, emotionally demanding, monotonous, and repetitive jobs, where the risk of developing neck and shoulder disorders is high. Additional explanations could be that the physical work environment is better suited for men than for women, so that chairs and desks are too high for shorter women and tools are too big and heavy for women's smaller hands. Also the work organization is usually designed by men and may be less adequate for women. Lack of unwinding after work may further contribute to the higher health risks for women.

Concluding Remarks

Numerous studies suggest an important role not only for physical but also for psychosocial stress at work in major health problems, and several plausible psychobiological mechanisms have been proposed. Two important neuroendocrine stress systems – the sympathetic adrenal medullary and the HPA systems – seem to play an important role for these relationships.

Information on psychobiological pathways between psychosocial stress at work and symptoms and illness is of great importance for the possibility of preventing and treating work-related ill health. Such knowledge could also help individual workers to better understand their own health problems and find ways of breaking vicious circles that form between work stress, behavior, and symptoms. It may also help to create a more understanding attitude in general toward people who, under stressful work conditions, have developed somatic symptoms. However, most important is that knowledge about these mechanisms is a strong incentive for implementing structural and organizational changes at work in order to protect people from mental and physical disorders. The psychobiological parameters may be used as warning signals but may also offer new possibilities for evaluating improvement of work conditions before they are reflected in somatic illness patterns, which usually (but not always) take a long time to develop. With regard to the enormous costs and suffering associated with musculoskeletal disorders, actions leading even to a moderate or small reduction of these problems would be extremely cost efficient.

The new forms of work that have become possible through the development of communication technology and that allow maximal flexibility and mobility have the potential for beneficial as well as harmful stress effects. The balance between negative and positive consequences of this type of boundaryless work is likely to be determined by factors such as the organization of work, the creation of new boundaries between different roles, and individual factors. An adequate balance between activation (energy mobilization) and rest and recovery (e.g., recreation, healing, growth) is necessary for long-term health.

Acknowledgment

Financial support has been obtained from the Swedish Council for Work Life Research, the Swedish Research Council, and the Bank of Sweden Tercentenary Foundation.

See Also the Following Articles

Allostasis and Allostatic Load; Economic Factors and Stress; Education Levels and Stress; Gender and Stress; Industrialized Societies; Psychosocial Factors and Stress; Type A Personality, Type B Personality; Work–Family Balance.

Further Reading

Bongers, P. M., de Winter, C. R., Kompier, M. A. J. and Hildebrandt, V. H. (1993). Psychosocial factors at work and musculoskeletal disease. *Scandinavian Journal of Work Environment and Health* **19**, 297–312.

Bosma, H., Peter, R., Siegrist, J. and Marmot, M. (1998). Alternative job stress models and risk of coronary heart disease. The effort-reward imbalance model and the job strain model. *American Journal of Public Health* **88**, 68–74.

Frankenhaeuser, M., Lundberg, U. and Chesney, M. (eds.) (1991). *Women, work and stress* New York: Plenum.

House, J. S., Landis, K. R. and Umberson, D. (1988). Social relationships and health. *Science* **241**, 540–545.

Karasek, R. A. and Theorell, T. (1990). *Healthy work. Stress, productivity, and the reconstruction of working life*. New York: Basic Books.

Lundberg, U. (1996). The influence of paid and unpaid work on psychophysiological stress responses of men and women. *Journal of Occupational Health and Psychology* **1**, 117–130.

Lundberg, U. and Hellström, B. (2002). Workload and morning salivary cortisol in women. *Work and Stress* **16**, 356–363.

Lundberg, U. and Johansson, G. (2000). Stress and health risks in repetitive work and supervisory monitoring work. In: Backs, R. & Boucsein, W. (eds.) *Engineering psychophysiology: issues and applications*, pp. 339–359, Hillsdale, NJ: Lawrence Erlbaum Associates.

Lundberg, U., Mårdberg, B. and Frankenhaeuser, M. (1994). The total workload of male and female white collar workers as related to occupational level, number of and age. *Scandinavian Journal of Psychology* **35**, 315–327.

Marmot, M. G., Davey Smith, G., Stansfeld, S., Patel, C., North, F., Head, J., et al. (1991). Health inequalities among British civil servants: the Whitehall II study. *Lancet* **337**, 1387–1393.

McEwen, B. S. and Lasley, E. N. (2002). *The end of stress as we know it*. Washington DC: Joseph Henry Press.

Ryff, C. D. and Singer, B. (1998). The contours of positive human health. *Psychological Inquiry* **9**, 1–28.

Sapolsky, R. M. (1998). *Why zebras don't get ulcers. An update guide to stress, stress-related diseases, and coping*. New York: W.H. Freeman and Co.

Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal of Occupational Health and Psychology* **1**, 27–41.

Williams, R. B., Barefoot, J. C. and Shekelle, R. B. (1985). The health consequences of hostility. In: Chesney, M. A. & Rosenman, R. H. (eds.) *Anger and hostility in cardiovascular and behavioral disorders*, pp 173–185. New York: McGraw-Hill.